Working Paper 98-21
Statistics and Econometrics Series 06
July 1998

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34)91 624-9849

MODELIZACIÓN DE SERIES TEMPORALES FINANCIERAS. UNA RECOPILACIÓN.

Begoña Font*

Abstract

This work reviews the main univariate and multivariate models proposed in the literature to represent the second moments dynamics. The paper starts with a description of stylized facts of financial time series and follows with a comparative study of different models available for modelling these characteristics. The main statistical properties of ARCH and stochastic volatility models are analyzed, as well as, the state space models with dynamics in the variance, Markow switching-variance models and microstructure models. Bayesian contributions to the field are also reviewed.

Keywords:
ARCH; Bayesian estimation; financial time series; heterocedasticity; Markow switching-variance models; Monte Carlo Markow Chains; microstructure models; state space models; stochastic volatility.

*Font, Universitat de Valencia, Departamento de Economía Financiera y Matemática, España. E-mail: Maria.B.Font@uv.es. Este artículo se desarrolló durante una estancia de investigación en el Departamento de Estadística y Econometría de la Universidad Carlos III de Madrid. La autora aprovecha estas líneas para agradecer la amable acogida dispensada por el Departamento y en especial por el Dr. Daniel Peña. El trabajo ha sido parcialmente financiado por el Proyecto de la DGES PB96-0776 y la beca POST98-EJ-16-129 de la Consellería de Cultura, Educació y Ciència dentro del Plan Valenciano de Ciencia y Tecnología.

1. Introducción

Las series financieras tienen características singulares en relación a otras series macroeconómicas, como son: la mayor frecuencia en la observación de los datos, semanal, diaria o incluso minuto a minuto, y sobre todo la presencia de heterocedásticidad, que hace inadecuados los modelos desarrollados para series estacionarias. Estas características diferenciadoras han propiciado numerosos trabajos en las áreas de econometría y economía financiera desde los años 70. Así, en el área financiera las evidencias sobre estructuras de volatilidades implícitas[1] ha conducido a la formulación de modelos de valoración continuos con volatilidad no constante[2], y a su discretización y estimación como series temporales[3]. Los modelos de volatilidad estocástica que estudiaremos en este artículo pueden obtenerse como resultado de estas discretizaciones. Por otra parte, en el área econométrica, el modelo ARCH propuesto por Engle(1982) para modelización de series heterocedásticas, por sus atractivas propiedades estadísticas, que le permiten captar las características más habituales de estas series y por la sencillez de su estimación, marca el punto de partida de estos modelos y sus generalizaciones que tienen también su reflejo en el área financiera[4].

El objetivo de este artículo, es la recopilación y revisión de los modelos estudiados por la literatura para la modelización de los hechos estilizados que presentan las series temporales financieras. Otros artículos de recopilación dedicados a la revisión de estos modelos son: Bollerslev, Chou and Kroner(1992), Bollerslev, Engle and Nelson(1995), Bera and Higgins(1993) y Palm(1996) centrados en los modelos ARCH, Taylor(1994) y Ghysels, Harvey and Renault(1996) en los modelos de volatilidad estocástica y, Ruiz(1994) y Shephard(1996) en un análisis comparativo entre las modelizaciones ARCH y de volatilidad estocástica. La literatura de recopilación y revisión de estos modelos (a la que podemos añadir algunos libros monográficos, citemos por ejemplo, Gouriéroux(1997) para los modelos ARCH y el clásico Taylor(1986) para modelos discretos de volatilidad estocástica) es bastante extensa. La recopilación que presentamos a continuación revisa los modelos ARCH y de volatilidad estocástica discretos y sus propiedades, presenta las últimas aportaciones en modelización y estimación realizadas en estas dos líneas de trabajo y, adicionalmente, revisa otros modelos también de aplicación en series temporales financieras pero no analizados en las revisiones antes citadas. Otra diferencia que presenta esta recopilación es un mayor hincapié en la revisión de la estimación Bayesiana de todos estos modelos.

El artículo se divide en 8 secciones incluida la presente introducción. En la sección 2 se analizan las características habituales que presentan las series temporales financieras y algunas ideas básicas para su modelización. En las secciones 3 y 4 se revisan, respectivamente, los modelos ARCH y de volatilidad estocástica discretos: modelos básicos, propiedades estadísticas, estima-

[1]La volatilidad implícita es la que se obtiene al despejar como incognita la volatilidad en la fórmula de Black and Scholes(1973) para valoración de opciones.

[2]Como ejemplos de modelos de valoración continuos tenemos los analizados en: Hull and White(1987), Scott(1987), Stein and Stein(1991) y Heston(1993).

[3]Ver Clarke(1973).

[4]Como referncias, ver por ejemplo, Nelson(1990a) y Duan(1997) que obtienen los procesos de difusión límite de algunos modelos ARCH, y el modelo de valoración GARCH para opciones propuesto en Duan(1995).

ción y generalizaciones. En la sección 5 se establecen comparaciones entre las modelizaciones ARCH y de volatilidad estocástica. En la sección 6 se revisan otras modelizaciones sobre series temporales financieras: modelos en el espacio de estados, modelos con cambios de régimen de la volatilidad y modelos de microestructura de mercados financieros. En la sección 7 se analiza la estimación Bayesiana de los modelos presentados explicando las ventajas de este procedimiento de estimación. Y en la sección 8 se incluyen conclusiones y últimos comentarios.

2. Hechos estilizados de las series teporales financieras e ideas básicas para su modelización

Las series temporales financieras se caracterizan por los siguientes rasgos comunes conocidos como hechos estilizados de las series temporales financieras.

a.-) *Nula o escasa estructura dinámica en la media.* Habitualmente, la serie en sí misma o en su primera diferencia presenta media constante a lo largo del tiempo o una estructura dinámica que puede describirse mediante un modelo ARMA con parámetros pequeños. En consecuencia, la serie original, transformada o los residuos de la serie en el caso de un ajuste ARMA constituyen series de media cero y no autocorreladas (estadísticos Box-Ljung no significativos).

b.-) *Estructura dinámica en la varianza.* Las series originales y sus transformaciones presentan una estructura dinámica para la varianza (son heterocedásticas), que se manifiesta en autocorrelaciones muestreales para los cuadrados significativamente no nulas. Esta estructura presenta algunas características que analizaremos en los próximos puntos bajo los epígrafes de: agrupación de la volatilidad (*volatility clustering*), efectos simétricos y de apalancamiento (*leverage effects*) y persistencia (*long-memory*).

c.-) *Distribuciones leptocúrticas (thick tails).* El estadístico de curtosis para la series financieras supera habitualmente el valor de 3.

d.-) *Agrupamiento de la volatilidad (volatility clustering).* Habitualmente, cambios pequeños vienen seguidos de cambios pequeños y cambios grandes de cambios grandes. Como consecuencia, detectamos en los gráficos de estas series zonas de volatilidad baja y alta.

e.-) *Efectos simétricos y de apalancamiento.* En las series sobre tipos de interés o de cambio, las variaciones de precio dan lugar a variaciones simétricas de la volatilidad y en consecuencia a distribuciones simétricas. En cambio, en las series sobre precios de activos o derivados se suele observar una respuesta asimétrica de la volatilidad con evidencias empíricas de correlaciones negativas entre rentabilidad y volatilidad y sobre mayores incrementos de la volatilidad cuando la rentabilidad es inferior a la esperada que cuando la supera. También hay evidencias de correcciones posteriores que realiza el mercado cuando considera que se ha sobrevalorado el riesgo o que el aumento de rentabilidad del activo no viene acompañado de aumento de riesgo (correlaciones positivas entre rentabilidad y volatilidad).

f.-) *Persistencia de la volatilidad (long-memory).* La evidencia empírica obtenida mediante el ajuste de modelos ARCH y SV indica que los efectos de cambio en la volatilidad tienen una alta persistencia en el tiempo.

g.-) *Cambios de régimen en la volatilidad.* Estos cambios pueden deberse a cambios en el ciclo económico, cambios estructurales del mercado, circunstancias particulares del activo, etc.

h.-) *Efectos debidos a la incorporación de información en el mercado.* Estos efectos adquieren mayor importancia al aumentar la frecuencia de los datos de la serie (por ejemplo observaciones en los momentos de las transacciones). La literatura en series diarias reconoce los efectos estacionales por cierre de mercado en el viernes y lunes y por vacaciones, efectos debidos a predicciones sobre algún hecho que pueda modificar el valor del activo, por ejemplo el anuncio de dividendos, etc. Las modelizaciones de microestructura de mercados basadas en series de alta frecuencia consideran efectos en la volatilidad por el número y tipo de agentes que participan en el mercado, tiempo entre transacciones, volumen de transacciones, etc.

i.-) *Efectos comunes.* Hay evidencias empíricas de grupos de series (sobre tipos o activos) con comportamientos en lo que se refiere a volatilidad similares. Es más, la mayor parte de la teoría financiera considera modelos multivariantes para la valoración de activos, formación de cartera óptima y estructura de los tipos de interés.

Las evidencias (a) a (g) e (i) aparecen ampliamente documentadas por la literatura financiera, ver por ejemplo: Mandelbrot(1963, 67), Fama(1963, 65), Black(1976) y Christie(1982) entre otros. Las evidencias del tipo (h) aparecen comentadas y revisadas en el artículo de Hasbrouck(1991) y el libro de O'Hara(1995).

Los hechos estilizados de las series temporales financieras sirven de base para la introducción de las siguientes ideas básicas para su modelización. Si denotamos por y_t al valor de la serie en estudio en el instante t, $\mathbf{x}_t$ a un vector de características exógenas en el instante t y por $\mathbf{y}^{(t-1)} = (y_0, y_1, \ldots, y_{t-1})'$ y $\mathbf{x}^{(t-1)} = (\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{t-1})$ a los vectores y matriz con la información disponible en t sobre y y $\mathbf{x}$ respectivamente[5], de acuerdo con (a) y (h) (introduciendo los efectos sobre entrada de información como variables exógenas), podemos considerar el modelo:

$$y_t = g(\mathbf{y}^{(t-1)}; \mathbf{x}^{(t-1)}) + \varepsilon_t, \tag{2.1}$$

con $g(.)$ una función de $\mathbf{y}^{(t-1)}$ y $\mathbf{x}^{(t-1)}$. Para modelizar esta estructura la literatura propone un modelo ARMA, un modelo de regresión sobre variables exógenas y/u observaciones retardadas, o un modelo de regresión que incluya como variable explicativa una función de la volatilidad del proceso[6]. Otros autores sugieren otras estrategias dinámicas combinadas, así, Pagan and

[5] y_0 (que puede ser escalar, vector o no existir) representa la información sobre nuestra serie disponible en el instante 0, analogamente, $\mathbf{x}_0$ (que puede ser vector, matriz o no existir) representa características exógenas conocidas en el instante 0.

[6] Como referencias destaquemos: los modelos ARMA-ARCH estudiados por Weiss(1984), los modelos de regresión ARCH de Engle(1982) y Geweke(1988a), los modelos de regresión SV de Harvey and Shephard(1993c)

Schwert(1990) propone eliminar los efectos día de la semana (regresando sobre variables dummies que indican a qué día de la semana pertenece cada observación entre los 5 laborables) y aplicar sobre el residuo resultante un AR(5). Eliminada la estructura dinámica de la media a través de (2.1) se considera la modelización del ruido ε_t. Para modelizar el ruido, la teoría, habitualmente, asume que:

$$\varepsilon_t = z_t\sqrt{h_t}, \tag{2.2}$$

donde las variables aleatorias $\{z_t\}$ y $\{h_t\}$ son independientes contemporaneamente y en sus retardos, $\{z_t\}$ son variables aleatorias independientes e identicamente distribuidas de media 0 y varianza 1 ($z_t \sim \mathrm{IID}(0,1)$) y las variables $\{h_t\}$ pueden ser observables en t ($h_t = h_t(\boldsymbol{\varepsilon}^{(t-1)}, \mathbf{h}^{(t-1)})$), por ejemplo, un modelo ARCH, o no observable en t, por ejemplo, un modelo SV. [7]

Notemos que de (2.2) se tiene que ε_t es una diferencia martingala con primeros momentos marginales dados para t y $s \neq 0$ por:

$$\begin{array}{rcl} \mathrm{E}(\varepsilon_t) & = & 0, \\ \mathrm{V}(\varepsilon_t) & = & \mathrm{E}(h_t), \\ \mathrm{C}(\varepsilon_t, \varepsilon_{t-s}) & = & 0, \end{array} \tag{2.3}$$

quedando la varianza de la serie como función de h_t, de ahí que la literatura defina $\sqrt{h_t}$ como la volatilidad de la serie. Las distintas modelizaciones que asumamos sobre h_t junto con las hipótesis distribucionales sobre z_t justificaran los distintos hechos estilizados observados en las series temporales financieras (b) a (g). Los efectos (i) se recogen a través de la generalización multivariante de estos argumentos.

3. Modelos ARCH (Autoregressive Conditional Heterocedasticity)

Los modelos ARCH (Autoregressive Conditional Heterocedasticity) son introducidos por Engle(1982) para la modelización de la inflacción[8]. Estos modelos asumen que la volatilidad de la serie es observada en t y es función de los retardos de ε y de h, de modo que, a partir de (2.2) los primeros momentos condicionales vienen dados por:

$$\begin{array}{rcl} \mathrm{E}(\varepsilon_t|\boldsymbol{\varepsilon}^{(t-1)}) & = & 0, \\ \mathrm{V}(\varepsilon_t|\boldsymbol{\varepsilon}^{(t-1)}) & = & h_t = h_t(\boldsymbol{\varepsilon}^{(t-1)}, \mathbf{h}^{(t-1)}). \end{array} \tag{3.1}$$

Este supuesto de observabilidad de la volatilidad permite la fácil deducción de la función de verosimilitud y en consecuencia una estimación, predicción y diagnosis del modelo "fácil" desde

y dentro del último grupo los modelos ARCH-M de Engle, Lilien and Robbins(1987) y Bollerslev, Engle and Wooldridge(1988) y el modelo SV-M de Pitt and Shephard(1995).

[7] Todas estas notaciones y las ecuaciones (2.1) y (2.2) pueden extenderse al caso de análisis multivariante de varias series. Ver subsecciones 3.4, 4.4 y el modelo (6.2).

[8] La serie macroeconómica de la inflacción presenta algunas características similares a las de las series temporales financieras.

un punto de vista clásico, y proporciona un procedimiento sencillo (a través de la especificación de la función $h_t(.)$) para representar las características de agrupamiento, efectos simétricos y de apalancamiento de la volatilidad. Los costes de estas ventajas son, en principio, la hipótesis de observabilidad de la volatilidad y las restricciones a imponer sobre los parámetros de la función $h_t(.)$ para asegurar la positividad de la varianza condicional.

Los modelos ARCH, que revisamos a continuación, han sido objeto de numerosos estudios que abarcan diversas generalizaciones, estudios sobre las propiedades, estimación de parámetros, predicción y numerosas aplicaciones financieras.

3.1. Modelos ARCH básicos

La literatura sobre modelos ARCH es muy amplia y comprende muchas generalizaciones de los modelos iniciales propuestos por Engle(1982), Bollerslev(1986) y Engle and Bollerslev(1986) para captar aspectos específicos de las series temporales financieras. En esta subsección presentaremos los modelos básicos y describiremos sus propiedades, dejando para la subsección 3.4 la descripción de algunas de sus generalizaciones más interesantes.

Los modelos básicos son:

a.-) Modelo ARCH(q):

$$\left.\begin{array}{rcl} \varepsilon_t & = & z_t\sqrt{h_t} \\ z_t & \sim & \mathrm{IID}(0,1) \\ h_t & = & \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \\ & & \omega > 0,\ \alpha_i \geq 0,\ \forall i \end{array}\right\}. \tag{3.2}$$

[Engle(1982)]

b.-) Modelo GARCH(p,q):

$$\left.\begin{array}{rcl} \varepsilon_t & = & z_t\sqrt{h_t} \\ z_t & \sim & \mathrm{IID}(0,1) \\ h_t & = & \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \\ & & \omega > 0,\ \ \alpha_i \geq 0,\ \ \beta_i \geq 0,\ \ \forall i \end{array}\right\}. \tag{3.3}$$

[Bollerslev(1986)]

c.-) Modelo IGARCH(p,q):

$$\left.\begin{array}{rcl} \varepsilon_t & = & z_t\sqrt{h_t} \\ z_t & \sim & \mathrm{IID}(0,1) \\ h_t & = & \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \\ & & \omega > 0,\ \ \alpha_i \geq 0,\ \ \beta_i \geq 0,\ \ \forall i \\ & & \sum_{i=1}^{q} \alpha_i + \sum_{i=1}^{p} \beta_i = 1 \end{array}\right\}. \tag{3.4}$$

[Engle and Bollerlev(1986)]

d.-) Modelo log-ARCH(q):

$$\left.\begin{array}{lcl} \varepsilon_t & = & z_t\sqrt{h_t} \\ z_t & \sim & \mathrm{IID}(0,1) \\ \log h_t & = & \gamma_0 + \sum_{i=1}^{q} \gamma_i \log \varepsilon_{t-i}^2 \end{array}\right\}. \qquad (3.5)$$

[Geweke(1986)]

Del análisis del modelo (3.2) de Engle(1982) (el proceso más sencillo) se desprende que presenta propiedades estadísticas adecuadas para modelizar los hechos estilizados de ausencia de estructura dinámica de la media (a), estructura dinámica de la varianza (b), distribución leptocúrtica (c), agrupamiento de volatilidad (d) y efectos simétricos (e.1). Así, las características (a), (b), (d) y (e.1) se deducen de las expresiones de los momentos condicionales para el modelo (ver (3.1) y (3.2)), y la propiedad (c) y la simetría de la distribución marginal de ε_t se obtiene de la aplicación sobre (3.2) de la ley de expectativas reiteradas y los supuestos distribucionales sobre z_t. En particular, para el modelo ARCH normal se tiene que la distribución marginal es simétrica, de varianza finita y constante (si $\sum_{i=1}^{q} \alpha_i < 1$) y leptocúrtica para $q = 1$ (bajo restricciones adicionales sobre los parámetros para asegurar la existencia de los momentos). Las expresiones sobre los primeros momentos marginales para el modelo ARCH(q) y el coeficiente de curtosis para el modelo ARCH(1) normal se obtienen de (3.6) y (3.7) tomando $\beta_i = 0$, $i = 1, \ldots, p$.

Los parámetros α determinan la heterocedasticidad del proceso. La aplicación empírica de estos modelos en series financieras da lugar a ajustes con órdenes q bastante altos ($q \approx 6$ ó 7) y parámetros α cercanos a cero. Estos órdenes aumentan considerablemente el número de parámetros α a estimar (que además deben cumplir restricciones para garantizar la positividad de la varianza y restricciones adicionales si queremos asegurar la estacionariedad en sentido amplio y existencia de la varianza marginal (ver subsección 3.2). Engle(1982) propone la reparametrización $\alpha_i = \phi(q+1-i)/(q(q+1))$, $i = 1, \ldots, q$, que reduce a dos (ω, ϕ) los parámetros a estimar, y posteriormente, Bollerslev(1986) y Taylor(1986) en trabajos independientes proponen el modelo GARCH.

Los modelos GARCH (ver (3.3)), generalizan los modelos ARCH y mantienen sus propiedades estadísticas de interés en relación a las modelización de los hechos estilizados (a), (b), (c), (d) y (e.1), y proporcionan buenos ajustes con p y q pequeños (la mayoría de las series temporales financieras pueden modelizarse correctamente con un GARCH(1,1)). Bollerslev(1986) proporciona la justificación teórica de esta última afirmación expresando los procesos GARCH(p,q) como un ARCH(∞). Otra propiedad importante de los modelos GARCH, de interés en el área financiera, es que son una aproximación a procesos de difusión. Así, Nelson(1990a) prueba la convergencia del modelo GARCH(1,1) con errores condicionales normales a un proceso de difusión continuo con distribuciones estacionarias no condicionadas t. Otras referencias sobre resultados de este tipo son: Nelson(1992, 95, 96) y Nelson and Foster(1994, 95) y Duan(1997).

En relación a las propiedades marginales de los GARCH, señalemos que, bajo normalidad, la distribución marginal es simétrica, de varianza constante (si $\sum_{i=1}^{q}\alpha_i + \sum_{i=1}^{p}\beta_i < 1$) y la covarianza serial cero. Además en el caso del proceso GARCH(1,1) normal tenemos que la función de autocorrelación de los cuadrados (que existe si $\alpha_1 + \beta_1 < 1$) es una función de las potencias de $\alpha_1 + \beta_1$. Por lo que, generalizando podemos interpretar $\sum_{i=1}^{q}\alpha_i + \sum_{i=1}^{p}\beta_i$ como una medida de la persistencia de los efectos de volatilidad (y como empíricamente los α son pequeños asociaremos en concreto el efecto persistencia a los parámetros β). La estimación de series temporales financieras con modelos GARCH(1,1) obtiene valores para $\alpha_1 + \beta_1$ habitualmente próximos a 1, lo que sugiere el hecho estilizado de alta persistencia de volatilidad (f).

Engle and Bollerslev(1986) proponen para modelizar la persistencia de volatilidad el modelo IGARCH (Integrado GARCH). Los procesos IGARCH(p,q) son adecuados para modelizar los hechos estilizados (a) a (e.1), y el de persistencia (f), tienen varianza marginal infinita, son estrictamente estacionarios y empiricamente suelen dar lugar a buenos ajustes.

Notemos que, en los modelos ARCH, GARCH e IGARCH que acabamos de estudiar se imponen restricciones para asegurar la positividad de la varianza condicional. Así, las restricciones introducidas en las ecuaciones (3.2), (3.3) y (3.4) son suficientes para esta positividad aunque no necesarias (Nelson and Cao(1992) proponen condiciones más suaves para esta positividad), y la condición $\omega > 0$ es necesaria para la no degeneración de los procesos ARCH(1), GARCH(1,1) estacionarios e IGARCH(1,1) (ver Nelson(1990b)).

Geweke(1986) propone un modelo en el que no es necesario establecer restricciones sobre los parámetros para asegurar la positividad de la varianza condicional. Este modelo conocido por log-ARCH(q) (ver (3.5)) también captura los hechos estilizados (a), (b), (c) y (e.1). Su mayor diferencia con los anteriores es que el efecto sobre la varianza condicional es multiplicativo mientras que en los otros tres era aditivo. El estudio empírico de Higgins and Bera(1992) argumenta mejores resultados de modelización con este modelo que los obtenidos con el modelo ARCH.

Para terminar, señalemos que al estudiar las propiedades marginales de estos modelos nos hemos restringido a modelos normales (con variables z_t normales). También se han estudiado, con el objetivo de modelizar mejor el efecto leptocúrtico, modelos ARCH con errores t de Student, mixturas de normal-Poisson, mixturas normal-lognormal y exponenciales y, como alternativa las distribuciones exponenciales generales (GED) que incluyen como caso particular el normal[9]. Como referencias indiquemos: Harvey(1981), Bollerslev(1987), Jorion(1988), Hsieh(1989) y Nelson(1991).

3.2. Propiedades estadísticas

Las propiedades estadísticas que indicamos a continuación se refieren a los modelos ARCH con errores condicionales normales[10].

[9] En referencia a las distribuciones GED citemos: Box y Tiao(73)

[10] La hipótesis de normalidad no es necesaria para los resultados de estacionariedad amplia y estricta ni para la obtención de los primeros dos momentos (ecuación (3.6)).

Momentos Marginales

Bollerslev(1986) demuestra que bajo hipótesis de normalidad, los procesos GARCH(p,q) con $\sum_{i=1}^{q}\alpha_i + \sum_{i=1}^{p}\beta_i < 1$ son estacionarios en sentido amplio y sus primeros momentos marginales vienen dados para todo t y $s \neq 0$ por:

$$\begin{array}{lcl} \mathrm{E}(\varepsilon_t) & = & 0, \\ \mathrm{V}(\varepsilon_t) & = & \frac{\alpha_0}{1-\sum_{i=1}^{q}\alpha_i-\sum_{i=1}^{p}\beta_i}, \\ \mathrm{C}(\varepsilon_t,\varepsilon_{t-s}) & = & 0. \end{array} \tag{3.6}$$

También obtiene fórmulas recursivas para el cálculo de los primeros momentos del modelo GARCH(1,1). En particular, obtiene que todos los momentos de orden impar son cero y si $3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1$ el coeficiente de curtosis es finito (expresión (3.7)) y mayor que 3

$$\kappa_\varepsilon = 3 + \frac{6\alpha_1^2}{1 - 3\alpha_1^2 - 2\alpha_1\beta_1 - \beta_1^2}. \tag{3.7}$$

Como consecuencias, los procesos GARCH(p,q) estacionarios en sentido amplio son diferencias martingala con varianza finita constante y los GARCH(1,1) presentan distribuciones para ε simétricas y leptocúrticas. Estas mismas observaciones y las expresiones anteriores de los momentos marginales son trasladables a los modelos ARCH(q) y ARCH(1) respectivamente, tomando en todas las expresiones $\beta_i = 0$, $i = 1,\ldots,p$ (los resultados para los modelos ARCH(q) fueron probados por Engle(1982)).

Por otra parte, Nelson(1990b) demuestra que los procesos ARCH no son necesariamente estacionarios en sentido estricto aunque lo sean en sentido amplio. Para ver más detalles sobre las condiciones que debe cumplir un proceso GARCH(1,1) para ser estacionario en sentido estricto el lector interesado puede referirse a Nelson(1990b), y para ver otras condiciones de estacionariedad (desde un punto de vista Bayes) a Kleibergen and Van Dijk(1993).

El modelo IGARCH(p,q) no tiene varianza marginal finita y no es estacionario en sentido amplio. Aunque, Nelson(1990b) demuestra que el proceso IGARCH(1,1) es estacionario en sentido estricto y ergódico, estos resultados son generalizados para el proceso IGARCH(p,q) por Bougerol and Picard(1992).

Función de autocorrelación de los cuadrados (fac)

Bollerslev(1988) demuestra que las autocorrelaciones de ε_t^2 para un GARCH(1,1) normal vienen dadas por:

$$\rho_{\varepsilon_t^2}(s) = \frac{C(\varepsilon_t^2,\varepsilon_{t-s}^2)}{V(\varepsilon_t^2)} = \begin{cases} \frac{\alpha_1(1-\alpha_1\beta_1-\beta_1^2)}{1-2\alpha_1\beta_1-\beta_1^2}, & s = 1, \\ (\alpha_1+\beta_1)^{s-1}\rho_{\varepsilon_t^2}(1), & s > 1. \end{cases} \tag{3.8}$$

Agregación temporal

Los procesos ARCH son cerrados bajo agregación temporal en términos de los dos primeros momentos y tenemos expresiones para relacionar los parámetros de los procesos de alta y baja frecuencia de datos, pero no son cerrados desde el punto de vista distribucional (ver Drost and Nijman(1992) y Nijman and Sentana(1996)).

3.3. Estimación de parámetros

La estimación clásica habitual se basa en la obtención del estimador máximo verosímil de los parámetros (estimación ML) o el estimador cuasi-máximo verosímil (estimación QML[11]) en ausencia de normalidad condicional. La aplicación del procedimiento de estimación ML y QML se apoya en lo sencillo que es, en estos modelos, la obtención de la función de verosimilitud y/o primeros dos momentos condicionada a unas pocas observaciones iniciales. Para ejemplificar el proceso consideremos el modelo GARCH(p,q) normal[12], si condicionamos sobre las q observaciones iniciales $(\varepsilon_{1-q}, \varepsilon_{2-q}, \dots, \varepsilon_0)$ y sobre[13] $(h_{1-p}, h_{2-p}, \dots, h_0)$, tenemos desde (3.3) considerando normalidad y dada una muestra $(\varepsilon_1, \dots, \varepsilon_T)$ la siguiente función de log-veromilitud:

$$
\begin{aligned}
&\log f(\varepsilon_1, \dots, \varepsilon_T | \varepsilon_{1-q}, \varepsilon_{2-q}, \dots, \varepsilon_0, h_{1-p}, h_{2-p}, \dots, h_0, \phi) \\
&= \text{const.} - \sum_{t=1}^{T} \log(h_t)/2 - \sum_{t=1}^{T} \varepsilon_t^2/(2h_t) \\
&= \text{const.} + \sum_{t=1}^{T} l_t, \qquad (3.9)
\end{aligned}
$$

donde $\phi = (\omega, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)'$ es el vector de parámetros a estimar y $l_t = -\log(h_t)/2 - \varepsilon_t^2/(2h_t)$, $t = 1, \dots, T$. Para obtener los estimadores ML de los parámetros maximizaremos (3.9). Las condiciones de primer orden son:

$$
\begin{aligned}
\frac{\partial l_t}{\partial \phi} &= \frac{1}{2h_t} \frac{\partial h_t}{\partial \phi} \left(\frac{\varepsilon_t^2}{h_t} - 1 \right), \\
\frac{\partial^2 l_t}{\partial \phi \partial \phi'} &= \left(\frac{\varepsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial \phi'} \left(\frac{1}{2h_t} \frac{\partial h_t}{\partial \phi} \right) - \frac{1}{2h_t^2} \frac{\partial h_t}{\partial \phi} \frac{\partial h_t}{\partial \phi'} \frac{\varepsilon_t^2}{h_t},
\end{aligned} \qquad (3.10)
$$

donde $\partial h_t / \partial \phi = \mathbf{w}_t + \sum_{i=1}^{p} \beta_i \partial h_{t-i} / \partial \phi$, con $\mathbf{w}_t = (1, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2, h_{t-1}, \dots, h_{t-p})'$, $t = 1, \dots, T$. La obtención del máximo de la función de log-verosimilitud (3.9) puede realizarse empleando, por ejemplo, el algoritmo de Berndt, Hall, Hall and Hausman(1974). La única dificultad del método ML estriba en la obtención de estimadores máximo verosímiles restringidos a determinadas condiciones sobre los parámetros para asegurar positividad, estacionariedad y en general la existencia de momentos marginales de orden superior a dos. En la actualidad, para realizar las estimaciones ML y QML de modelos ARCH podemos hacer uso de algunos paquetes estadísticos (Splus, Gauss, Eviews o TSP).

[11] La estimación QML consiste en la maximización de una función de verosimilitud que se obtiene asumiendo que las condicionales son normales con media y varianza dadas en (3.1).

[12] Las ecuaciones (3.9) y (3.10) son asimismo válidas para la estimación ML del modelo ARCH(q) normal con parámetros $\phi = (\omega, \alpha_1, \dots, \alpha_q)'$ tomando $\partial h_t / \partial \phi = \mathbf{w}_t$, con $\mathbf{w}_t = (1, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2)'$, $t = 1, \dots, T$.

[13] Bollerslev(1986) propone tomar como valores de la volatilidad en los p instantes iniciales $h_t = \widehat{\sigma}^2 = \sum_{t=1}^{T} \varepsilon_t^2 / T$, $t = 1 - p, \dots, 0$. Esta sugerencia no es necesaria para la estimación del modelo ARCH(q).

Como referencias teóricas sobre la estimación ML citemos a Engle(1982) y Bollerslev(1986) para modelos ARCH(q) y GARCH(p,q) normales[14] respectivamente, y a Crowder(1976) y Lumsdaine(1992) en referencia a la consistencia asintótica del estimador ML. En ausencia de normalidad condicionada de los errores, Weiss(1986) para modelos ARCH y, Bollerslev and Wooldridge(1992) y Gouriéroux(1992) para los GARCH, prueban la consistencia y normalidad asintótica de los estimadores QML. Otros artículos de interés adicionales sobre estimación QML y propiedades asintóticas del estimador son: Lumsdaine(1996) y Lee and Hansen(1994).

Otros procedimientos clásicos aplicados pero en menor medida son: la estimación Bayesiana, el método de los momentos generalizado (GMM) y la estimación indirecta. Aplicando estimación Bayesiana, Geweke(1988ab, 89ab) obtiene distribuciones posteriores y predictivas para un modelo de regresión con proceso sobre los errores ARCH(q) normal con $\alpha_i = \phi(q+1-i)/(q(q+1))$, empleando una distribución inicial sobre los parámetros impropia y aplicando muestreo importante e integración de Monte Carlo. Y Müller and Pole(1997) y Bauwens and Lubrano(1998) obtienen estimaciones Bayes de los parámetros basadas en distribuciones posteriores haciendo uso de Gibbs Sampler. Los procedimientos Bayesianos de estimación se verán con más detalle en la subsección 7.1.

3.4. Otros modelos ARCH

La consideración de la volatilidad como observable facilita en gran medida la modelización de los hechos estilizados de las series temporales financieras y ha dado lugar a muchas modelizaciones para captar determinado hecho estilizado, presentaremos a continuación algunas de las generalizaciones clasificándolas en función del hecho estilizado que fundamenta su modelización.

Modelización de efectos asimétricos y de apalancamiento

Constituye probablemente el grupo de mayor interés de modelización ya que hay muchas series financieras (sobre todo de activos y derivados) en las que el signo de la variación produce una respuesta asimétrica en la volatilidad. Destaquemos los siguientes modelos:

a.-) Modelo EGARCH(p,q):

$$\left.\begin{array}{lcl} \varepsilon_t & = & z_t\sqrt{h_t} \\ z_t & \sim & \mathrm{GED}(0,1) \\ \log h_t & = & \gamma_0 + \sum_{i=1}^{q} \gamma_i r(z_{t-i}) + \sum_{i=1}^{p} \gamma_i^b \log h_{t-i} \end{array}\right\}, \qquad (3.11)$$

donde $r(x) = \lambda_0 x + \lambda_1(|x| - E|x|)$.

[Nelson(1991)]

b.-) Modelo NGARCH:

[14]La estimación ML también se ha aplicado para distribuciones condicionales no normales, ver por ejemplo, Bollerslev(1987) y Nelson(1991).

$$\left.\begin{array}{rcl} \varepsilon_t & = & z_t\sqrt{h_t} \\ z_t & \sim & \mathrm{N}(0,1) \\ h_t & = & \omega + \beta_1 h_{t-1} + \beta_2 h_{t-1}(z_{t-1}-c)^2 \\ & & \omega > 0,\ \beta_1 \geq 0,\ \beta_2 \geq 0 \end{array}\right\}, \tag{3.12}$$

donde c es un parámetro adicional que suele interpretarse como una prima por riesgo.

[Engle and Ng(1993)]

c.-) Volatility-Switching ARCH:

$$\left.\begin{array}{rcl} \varepsilon_t & = & z_t\sqrt{h_t} \\ z_t & \sim & \mathrm{N}(0,1) \\ h_t & = & \omega + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1} + s_{t-1}\nu_{t-1} \\ \nu_{t-1} & = & \delta_0\varepsilon_t^2 - \delta_1 h_t - \delta_2 \\ s_t & = & \begin{cases} 1 & \text{si} \quad \varepsilon_t > 0 \\ 0 & \text{si} \quad \varepsilon_t = 0 \\ -1 & \text{si} \quad \varepsilon_t < 0 \end{cases} \\ & & \omega > 0,\ \alpha_1 \geq 0,\ \beta_1 \geq 0 \end{array}\right\}. \tag{3.13}$$

[Fornari and Mele(1997)]

Los modelos EGARCH y NGARCH, el primero multiplicativo y el segundo sumativo, están motivados para modelizar correlaciones negativas asimétricas entre rentabilidad de activos y volatilidad. El modelo Volatility-Switching ARCH captura esos efectos asimétricos y también efectos de correlaciones positivas entre rentabilidad y volatilidad como resultado de correcciones del riesgo del activo en cuestión.

Duan(1997) presenta un modelo Augmented GARCH(p,q) que recoge efectos multiplicativos y sumativos sobre la volatilidad, tiene como casos particulares los procesos: GARCH, log-GARCH, EGARCH, NGARCH y TGARCH y puede emplearse como aproximación discreta de varios modelos continuos de valoración de activos con volatilidad no constante.

Efectos debidos a la incorporación de información al mercado

Presentamos a continuación dos modelos, el primero de ellos (3.14) modeliza los efectos por cierre de mercado, mediante variables dummy (aplicando una técnica fácilmente modificable para representar otros fenómenos de llegada de información al mercado predecibles). Y el segundo (3.15) modeliza el establecimiento de precios límite en la cotización diaria de activos y derivados (un cambio normativo).

a.-) Modelo GARCH(1,1) con efectos cierre de mercado:

$$\left.\begin{array}{rcl} y_t & = & b_0 + \sum_{i=1}^{5} b_i D_{ti} + \varepsilon_t \\ \varepsilon_t & = & z_t\sqrt{h_t} \\ z_t & \sim & \mathrm{GED}(0,1) \\ h_t & = & \omega_0(1-\alpha_1-\beta_1) + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1} \\ & + & [1-(\alpha_1+\beta_1)L][\sum_{i=1}^{5}\omega_i D_{ti} + (\omega_6+\omega_7 L)V_t] \\ & & \omega_0 > 0,\ \omega_1 \geq 0,\ \alpha_1 \geq 0,\ \beta_1 \geq 0 \end{array}\right\}, \tag{3.14}$$

donde L es el operador retardo, las variables día de la semana dummy, D_{ti}, $i = 1, \ldots, 5$ toman el valor 1 cuando la observación corresponde al día en cuestión y la variable dummy, V_t, toma el valor 1 cuando el mercado se cierra por una razón diferente a la de fin de semana.

[Baillie and Bollerslev(1989)]

b.-) Modelo Censored-GARCH(1,1):

$$\left.\begin{array}{rcl} r_t & = & \begin{cases} \bar{c}_t & \text{si } r_t^* + \mathrm{LO}_{t-1} \geq \bar{c}_t \\ r_t^* + \mathrm{LO}_{t-1} & \text{si } \underline{c}_t < r_t^* + \mathrm{LO}_{t-1} < \bar{c}_t \\ \underline{c}_t & \text{si } r_t^* + \mathrm{LO}_{t-1} \leq \underline{c}_t \end{cases} \\ r_t^* & = & \mu + z_t\sqrt{h_t} \\ z_t & \sim & \mathrm{GED}(0,1) \\ h_t & = & \omega + \alpha(r_{t-1}^* - \mu)^2 + \beta h_{t-1} \\ & & \omega > 0,\ \alpha \geq 0,\ \beta \geq 0 \end{array}\right\}, \tag{3.15}$$

con p_t y r_t el precio y la log-rentabilidad respectivamente del activo cotizados en t, p_t^* y r_t^* el precio y la log-rentabilidad resultado de la intersección entre oferta y demanda respectivamente del activo en t $\underline{c}_t = \log(1 - a/p_{t-1})$, $\bar{c}_t = \log(1 + a/p_{t-1})$, a mide la banda permitida por el mercado para la cotización ($p_t \in (p_{t-1} - a, p_{t-1} + a)$) y $\mathrm{LO}_{t-1} = \log(p_{t-1}^*/p_{t-1})$.

[Wei(1998)[15]]

Otros modelos para captar efectos de incorporación de información al mercado se obtienen modelizando una estructura sobre la media (ver ecuación (2.1), sus comentarios y referencias).

Efectos de volatilidad permanentes y transitorios

Engle and Lee(1992) proponen el modelo Component-GARCH(1,1) definido por:

$$\left.\begin{array}{rcl} \varepsilon_t & = & z_t\sqrt{h_t} \\ z_t & \sim & \mathrm{N}(0,1) \\ h_t & = & q_t + \alpha_1(\varepsilon_{t-1}^2 - q_t) + \beta_1(h_{t-1} - q_t) \\ q_t & = & \omega + \rho(q_{t-1} - \omega) + \phi(\varepsilon_{t-1}^2 - h_{t-1}) \\ & & \omega > 0,\ \alpha_1 \geq 0,\ \beta_1 \geq 0,\ \rho \geq 0,\ \phi \geq 0 \end{array}\right\}. \tag{3.16}$$

[15] El modelo de Wei(1998) emplea estimación Bayesiana.

Este modelo permite reversión de la volatilidad a un nivel q_t que varía en el tiempo. El componente transitorio de la volatilidad es $h_t - q_t$ que converge a cero en potencias de $\alpha_1 + \beta_1$ y el componente a largo plazo (permanente) es q_t que converge a la varianza marginal de un proceso GARCH(1,1) en potencias de ρ. Empiricamente, ρ suele ser próximo a uno y la convergencia al valor de largo plazo es lenta.

Efectos comunes: modelos ARCH multivariantes

Los modelos ARCH multivariantes aparecen en la literatura de modelización casi con simultaneidad a los modelos univariantes (ver Kraft and Engle(1982)) con el objetivo de recoger efectos comunes de varias series temporales financieras de un mismo tipo, poder analizar el problema de determinación de cartera eficiente, testear la hipótesis de la teoría APT, etc. Básicamente, si denotamos por N al número de series modelizadas conjuntamente podemos hablar de dos tipos de modelos:

a.-) Modelo GARCH(p,q) multivariante:

$$\left.\begin{array}{lcl} \varepsilon_t & = & \mathbf{H}_t^{1/2}\mathbf{z}_t \\ \mathbf{z}_t & \sim & \mathrm{IID}(\mathbf{0},\mathbf{I}) \\ \mathrm{vec}(\mathbf{H}_t) & = & \boldsymbol{\omega} + \sum_{i=1}^{q}\mathbf{A}_i\mathrm{vec}(\varepsilon_{t-i}\varepsilon_{t-i}') + \sum_{i=1}^{p}\mathbf{B}_i\mathrm{vec}(\mathbf{H}_{t-i}) \end{array}\right\}, \tag{3.17}$$

donde $\mathbf{H}_t$ es una matriz de volatilidad de dimensión $N \times N$, $\mathbf{z}_t$ es un N-vector, $\boldsymbol{\omega}$ es un $N(N+1)/2$-vector de parámetros, $\mathbf{A}_i$ y $\mathbf{B}_i$ son matrices cuadradas $N(N+1)/2$ de parámetros y vec(.) es un operador que actua sobre una matriz simétrica colocando en columna los elementos de la parte triangular inferior.

b.-) Modelo GARCH K-factorial:

$$\left.\begin{array}{lcl} \mathbf{y}_t & = & \mathbf{B}\mathbf{f}_t + \varepsilon_t \\ \varepsilon_t & \sim & \mathrm{IID}(\mathbf{0},\boldsymbol{\Sigma}_\varepsilon) \\ \mathbf{f}_t & \sim & \text{proceso GARCH multivariante} \end{array}\right\}, \tag{3.18}$$

donde $\mathbf{B}$ es una matriz $N \times K$ de parámetros, ε_t es un N-vector, $\mathbf{f}_t$ es un K-vector de factores y $\boldsymbol{\Sigma}_\varepsilon$ es la matriz de varianzas y covarianzas de ε_t.

Al modelo multivariante GARCH definido en (3.17) hay que añadirle para su correcta especificación restricciones sobre los parámetros $\boldsymbol{\omega}$, $\mathbf{A}$ y $\mathbf{B}$ para garantizar que la matriz de varianzas-covarianzas condicional $\mathbf{H}$ sea definida positiva. La determinación de condiciones para asegurar esta positividad no son sencillas, como referencias sobre trabajos en esta línea señalemos: Engle, Granger and Kraft(1984), Baba, Engle, Kraft and Kroner(1990)[16], y Engle and Kroner(1995). Otro problema al que se enfrenta el modelo multivariante (3.17) es el número de parámetros a estimar para resolver esta cuestión, la literatura plantea varias soluciones:

[16] Estos autores presentan una reparametrización de la ecuación sobre la matriz de varianzas-covarianzas condicional conocida como representación BEKK en la que no se necesitan introducir restricciones sobre los parámetros

- Bollerslev, Engle and Wooldridge(1988) sugieren tomar matrices **A** y **B** diagonales.
- Bollerslev(1990) aconseja asumir que la matriz de correlaciones condicional es constante.
- Diebold and Nerlove(1989) estudian la teoría de que sólo unos pocos factores (no observables) son responsables de los efectos comunes que presentan grupos de series financieras y sugieren un modelo 1-factorial ARCH.

Esta tercera solución es la que se plantea con el modelo K-factorial ARCH definido en (3.18). Este modelo al que hay que añadirle para su correcta especificación las restricciones necesarias para la identificación de manera única de sus parámetros, constituye una generalización del modelo de Diebold and Nerlove(1989) para K-factores[17]. Engle(1987) propone otra alternativa de modelización para un modelo factorial en la que los factores son combinación lineal de los residuales. Otras referencias de interés son: Engle, Ng and Rothschild(1990), Ng, Engle and Rothschild(1992) y King, Sentana and Wadhawani(1994).

Otros modelos

Otras generalizaciones de modelos ARCH estudiadas en la literatura con sus respectivas referencias son: el modelo GJR-GARCH de Glosten, Jagannathan and Runkle(1993), el modelo TGARCH de Zakoian(1994), el modelo Qualitative Threshold ARCH de Gouriéroux and Monfort(1992), el modelo Nolineal ARCH de Higgins and Bera(1992), el modelo FIGARCH de Baillie, Bollerslev and Mikkelsen(1993), el modelo Unobserved ARCH[18] de Harvey, Ruiz and Sentana(1992), el modelo Semi-parametric ARCH de Engle and González-Rivera(1990), etc.

4. Modelos discretos de volatilidad estocástica (SV)

Los modelos discretos de volatilidad estocástica (SV) constituyen la segunda aproximación habitual a la modelización de series financieras. Estos modelos, estudiados por Taylor(1986, 94) parten de la modelización (2.2) y consideran la variable h_t como no observada introduciendo una estructura dinámica sobre sus retardos. El proceso resultante tiene la ventaja de una interpretación financiera sencilla como resultado de la discretización de modelos de valoración continuos (ver Clarke(1973)) y la característica de separar las perturbaciones sobre el nivel y sobre la varianza pero el inconveniente de una mayor complejidad en la estimación de sus parámetros debida a la difícil evaluación de la función de verosimilitud. En los próximos subapartados discutiremos algunos modelos, propiedades estadísticas y técnicas de estimación.

4.1. Modelos SV básicos

Los modelos SV también modelizan adecuadamente los hechos estilizados de las series financieras pero han sido menos estudiados por la literatura que los modelos ARCH. (Ver en la sección 5

[17] Los modelos K-factoriales tienen propiedades teóricas interesantes (surgen de la teoría del APT) pero su estimación no es fácil y requiere de la aplicación de métodos no lineales, ver por ejemplo Lin(1992)

[18] El modelo Unobserved ARCH comparte características estructurales de los modelos en el espacio de estados y ARCH.

en estudio comparado de las dos modelizaciones.) Introducidos por Taylor(1986) asumen en sus modelizaciones una estructura sobre $\log h_t$ autoregresiva por lo que se conocen también como procesos ARV (Autoregressive Random Variance). Revisaremos a continuación dos modelos: un modelo estacionario y un modelo de raíz unitaria.

a.-) Modelo ARV(p):

$$\left.\begin{array}{lll} \varepsilon_t & = & z_t\sqrt{h_t} \\ \log h_t & = & \gamma_0 + \sum_{i=1}^{p} \gamma_i \log h_{t-i} + \sigma_v v_t \\ \begin{pmatrix} z_t \\ v_t \end{pmatrix} & \sim & \mathrm{IID}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \end{array}\right\}. \tag{4.1}$$

[Taylor(1986)]

b.-) Modelo ARV(1) con una raíz unitaria:

$$\left.\begin{array}{lll} \varepsilon_t & = & z_t\sqrt{h_t} \\ \log h_t & = & \log h_{t-1} + \sigma_v v_t \\ \begin{pmatrix} z_t \\ v_t \end{pmatrix} & \sim & \mathrm{IID}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \end{array}\right\}. \tag{4.2}$$

El modelo ARV permite modelizar (como sucedía con el modelo ARCH) los hechos estilizados de ausencia de estructura dinámica en la media (a), estructura dinámica en la varianza (b), distribución leptocúrtica (c), agrupación de la volatilidad (d) y efectos simétricos (e.1). Efectivamente, de (4.1) es fácil ver que ε_t es una diferencia martingala con lo que justificamos (a), la estructura dinámica de la varianza (b) se introduce a partir del proceso autoregresivo sobre $\log h_t$ que proporciona el efecto de agrupación (d). Los efectos simétricos se deducen de la incorrelación entre z_t y v_t. Aplicando la ley de expectativas iteradas y las propiedades distribucionales asumidas se obtienen las expresiones de los momentos marginales y curtosis para el modelo ARV(1) normal (ecuaciones (4.3) y (4.4)) de las que se deduce la simetría de la distribución ε_t (otro argumento sobre el efecto simétrico) y su característica leptocúrtica que en este caso no depende de los parámetros.

Como ya comentábamos al principio de la sección, el modelo separa los efectos aleatorios sobre la media (modelizables a través de z_t) y sobre la varianza (modelizables a través de v_t) lo que permite en principio una mayor flexibilidad al modelo que la presentada por el modelo ARCH. Otra característica diferenciadora respecto al modelo ARCH (y recordemos la variante log-ARCH) es la no necesidad de imponer restricciones sobre los parámetros para garantizar la positividad de la varianza condicional. Siguiendo con los parámetros, γ_0 se puede interpretar como un factor de escala de la volatilidad del proceso[19], los γ_i son parámetros que miden la persistencia de los efectos sobre la volatilidad (mirar la expresión de la fac de un ARV(1) normal en (4.5)) y σ_v^2 mide la dispersión de la volatilidad (la heterocedásticidad del proceso).

[19]Notemos que podemos eliminar el parametro γ_0 del proceso autoregresivo sobre $\log h_t$, haciendo $\varepsilon_t = \sigma_* z_t \sqrt{h_t}$ con $\sigma_* = \exp(0.5\gamma_0)$.

El ajuste empírico del modelo ARV(1) proporciona valores para γ_1 próximos a la unidad señalando una alta persistencia de la volatilidad (hecho estilizado (f)). El modelo ARV(1) con raíz unitaria proporciona la alternativa más sencilla para modelizar esta persistencia. El modelo (4.2) presenta características estadísticas similares a los modelos IGARCH(1,1) en lo que se refiere a la modelizacion de hechos estilizados y como áquel tampoco tiene varianza marginal finita pero presenta una característica diferenciadora, el término independiente no necesita ser estimado ni incluido en la ecuación del proceso.

Para terminar, indiquemos que algunos autores (por ejemplo, Taylor(1994)) distinguen entre modelos contemporáneos ($\varepsilon_t = z_t\sqrt{h_t}$) y modelos retardados ($\varepsilon_t = z_t\sqrt{h_{t-1}}$).

4.2. Propiedades estadísticas

Nuevamente, las propiedades estadísticas que se estudian a continuación se establecen para los modelos SV básicos con distribución condicional normal[20].

Momentos marginales

Taylor(1994) demuestra para el modelo ARV(1), bajo las hipótesis de normalidad y $|\gamma_1| < 1$, que el proceso $\{\varepsilon_t\}$ es estacionario en sentido amplio y estricto y existen todos los momentos marginales. Los primeros momentos marginales vienen dados para t y $s \neq 0$ por:

$$\begin{array}{lcl} \mathrm{E}(\varepsilon_t) & = & 0, \\ \mathrm{V}(\varepsilon_t) & = & \exp\left(\frac{\gamma_0}{1-\gamma_1} + \frac{\sigma_v^2}{2(1-\gamma_1^2)}\right), \\ \mathrm{C}(\varepsilon_t, \varepsilon_{t-s}) & = & 0. \end{array} \tag{4.3}$$

También obtiene para el mismo modelo que los momentos impares son cero, y que el coeficiente de curtosis (expresión (4.4)) es mayor que 3. En consecuencia, la distribución marginal de ε_t es simétrica y leptocúrtica.

$$\kappa_\varepsilon = 3\exp\left(\frac{\sigma_v^2}{1-\gamma_1^2}\right). \tag{4.4}$$

Notemos que, a diferencia de los modelos ARCH no se necesitan introducir restricciones sobre los parámetros adicionales a las debidas a la estacionariedad para asegurar la existencia de momentos de orden superior a 2. Notemos también que el coeficiente de curtosis es superior a 3 de manera independiente al valor de los parámetros.

Expresiones sobre los momentos de la distribución marginal de ε_t bajo otras hipótesis distribucionales pueden encontrarse en Nelson(1991).

Función de Autocorrelación de los cuadrados (fac)

Bajo hipótesis de normalidad se tiene (siguiendo Taylor(1986)) que para el modelo ARV(1):

[20] La hipótesis de normalidad no es necesaria para los resultados de estacionariedad amplia y estricta pero si se aplica para la obtención de los momentos, curtosis y fac (ecuaciones (4.3), (4.4) y (4.5)).

$$\rho_{\varepsilon_t^2}(s) = \frac{\exp(\sigma^2_{\log h_t}\gamma_1^s) - 1}{3\exp(\sigma^2_{\log h_t}) - 1} \approx \frac{\exp(\sigma^2_{\log h_t}) - 1}{3\exp(\sigma^2_{\log h_t}) - 1}\gamma_1^s, \tag{4.5}$$

donde $\sigma^2_{\log h_t} = \sigma_v^2/(1-\gamma_1^2)$.

Agregación temporal

En referencia a la propiedad de agregación temporal, Meddahi and Renault(1995) demuestra que los modelos SV retardados de la forma:

$$\left.\begin{array}{lcl} \varepsilon_{t+1} & = & z_{t+1}\sqrt{h_t} \\ h_t & = & \omega + \beta_1 h_{t-1} + \sigma_v v_t \\ E(z_{t+1}|z_s, v_s, s \leq t) & = & 0 \\ E(z_{t+1}^2|z_s, v_s, s \leq t) & = & 1 \\ E(v_t|z_s, v_s, s \leq t) & = & 0 \end{array}\right\}, \tag{4.6}$$

forman una clase cerrada bajo agregación temporal[21].

4.3. Estimación de parámetros

La estimación en los modelos SV no es fácil ya que, al contrario de lo que sucedía con los modelos ARCH, no está claro cómo hacer la evaluación la función de verosimilitud y en consecuencia hay diferentes procedimientos para realizarla. Podemos señalar a grandes rasgos tres procedimientos de estimación: el método de momentos generalizado (GMM), la estimación cuasi-verosímil (QML) y los métodos clásicos y Bayesianos basados en simulación por cadenas de Markov Monte Carlo (MCMC). Describiremos a continuación los métodos de estimación QML y GMM, aplazando la revisión del método MCMC en su vertiente Bayesiana a la subsección 7.2.

El método de los momentos fue el primero que se propuso para la estimación de los parámetros de modelos SV. Este procedimiento, muy aplicado en la literatura financiera, se apoya en la existencia de un grupo de momentos marginales de los modelos SV (que representen la información más relevante del modelo) con expresiones teóricas conocidas y en los estimadores muestrales de los mismos, y consiste en obtener las expresiones de los parámetros que hagan mínima una distancia entre el vector con las expresiones teóricas de los momentos y el vector con los estimadores muestrales. La elección del grupo de momentos marginales a considerar y la elección de la distancia dan lugar a distintas variantes del método. Para fijar ideas, si consideramos el modelo ARV(1) (ver (4.1)) normal estacionario ($|\gamma_1| < 1$), una muestra $(\varepsilon_1, \ldots, \varepsilon_T)$ y la aplicación del método GMM para la selección de momentos y distancia de Melino and Turnbull(1990), el estimador óptimo de $\boldsymbol{\phi} = (\gamma_0, \gamma_1, \sigma_v)'$ es aquél tal que:

$$\widehat{\boldsymbol{\phi}} = \arg\min_{\boldsymbol{\phi}} \mathbf{q}'_T(\boldsymbol{\phi})\widehat{\boldsymbol{\Sigma}}_T^{-1}\mathbf{q}_T(\boldsymbol{\phi}), \tag{4.7}$$

donde

[21] El modelo (4.6) pertenece a la familia de modelos SARV definida por Andersen(1994).

$$
\begin{array}{rcl}
\mathbf{q}_T(\phi) & = & \frac{1}{T}\sum_{t=1}^{T}\mathbf{f}_t(\phi), \\
\mathbf{f}_t(\phi) & = & \left(\begin{array}{c}
\varepsilon_t^2 - \mathrm{E}(\varepsilon_t^2) \\
\varepsilon_t^4 - \mathrm{E}(\varepsilon_t^4) \\
|\varepsilon_t| - \mathrm{E}(|\varepsilon_t|) \\
|\varepsilon_t^3| - \mathrm{E}(|\varepsilon_t^3|) \\
\varepsilon_t^2\varepsilon_{t-1}^2 - \mathrm{E}(\varepsilon_t^2\varepsilon_{t-1}^2) \\
\vdots \\
\varepsilon_t^2\varepsilon_{t-10}^2 - \mathrm{E}(\varepsilon_t^2\varepsilon_{t-10}^2) \\
|\varepsilon_t\varepsilon_{t-1}| - \mathrm{E}(|\varepsilon_t\varepsilon_{t-1}|) \\
\vdots \\
|\varepsilon_t\varepsilon_{t-10}| - \mathrm{E}(|\varepsilon_t\varepsilon_{t-10}|)
\end{array}\right),
\end{array}
\tag{4.8}
$$

y $\widehat{\mathbf{\Sigma}}_T$ es un estimador consistente de $\mathbf{\Sigma}_T = \mathrm{E}(\sum_{t=1}^{T}\sum_{t'=1}^{T}\mathbf{f}_t(\phi)\mathbf{f}_{t'}(\phi)/T)$ Todos los momentos marginales considerados en (4.8) para el modelo considerado y sus expresiones teóricas son conocidas (ver Jacquier, Polson and Rossi(1994)).

Como referencias sobre estimación GMM destaquemos los artículos de Melino and Turnbull(1990) y Andersen and Sorensen(1996) respecto a implementación y el de Hansen(1982) en referencia a las propiedades asintóticas del estimador. Este método es sencillo de aplicar y no se apoya en las especificaciones distribucionales (aunque éstas pueden influir en la obtención de las expresiones de los momentos condicionales y en su existencia) pero tiene algunas limitaciones. Si nos centramos en el modelo ARV(1) normal (el modelo SV más sencillo), la estimación GMM necesita asumir de entrada (para garantizar la existencia de los momentos) la estacionariedad del proceso (una restricción sobre los parámetros a estimar), no proporciona estimaciones de la volatilidad (de gran interés en la literatura financiera) y presenta problemas de eficiencia cuando el parámetro γ_1 es próximo a 1 (situación bastante habitual, recordemos el rasgo estilizado de persistencia). Ruiz(1994) compara la eficiencia del método GMM (basado en las especificaciones de momentos de Melino and Turnbull(1990)) y del método QLM (basado en el filtro de Kalman) a través de las desviaciones asintóticas de los estimadores.

La alternativa habitual de estimación es el método QLM propuesto por Nelson(1988) y desarrollado para modelos SV estacionarios (ARV(1) con distribución condicional normal y no normal), modelos SV con una raíz unitaria y modelos SV asimétricos (ver (4.11)) por Harvey and Shephard(1993a,b), Ruiz(1994) y Harvey and Shephard(1996)[22]. El método QLM consiste en transformar el proceso SV hasta obtener su expresión en el espacio lineal de los estados y aplicar sobre este modelo de forma recurrente el filtro de Kalman para obtener media y varianza de las distribuciones condicionales que se aproximan como normales para así obtener la función de cuasi-verosimilitud, y a partir de su maximización obtener los estimadores QLM de los parámetros[23]. Revisemos, para concretar ideas, la estimación del modelo ARV(1) normal. Notemos, en primer lugar, que el modelo ARV(1) normal puede escribirse tomando $h_t^b = \log h_t$

[22]Los artículos de Ruiz(1994) y Harvey and Shephard(1996) incluyen expresiones asintóticas del estimador QLM.

[23]La metodología a seguir para aplicar el filtro de Kalman puede revisarse en Harvey(1989).

como:

$$
\begin{aligned}
\varepsilon_t &= z_t \exp(h_t^b/2), \\
h_t^b &= \gamma_0 + \gamma_1 h_{t-1}^b + \sigma_v v_t,
\end{aligned} \tag{4.9}
$$

y desde (4.9) tomando logaritmos sobre los cuadrados de ε_t y teniendo en cuenta la normalidad condicional como:

$$
\begin{aligned}
\log \varepsilon_t^2 &= \mathrm{E}(\log z_t^2) + h_t^b + u_t, \\
h_t^b &= \gamma_0 + \gamma_1 h_{t-1}^b + \sigma_v v_t,
\end{aligned} \tag{4.10}
$$

donde $u_t = \log z_t^2 - \mathrm{E}(\log z_t^2)$ es un ruido blanco no normal con varianza $V(u_t) = \pi^2/2$ y $\mathrm{E}(\log z_t^2) = -1.27$. La estimación QML considera el modelo (4.10), un modelo en el espacio de estados, y tratando el error u_t como un ruido blanco normal aplica el filtro de Kalman para obtener los dos primeros momentos de las distribuciones condicionales. A partir de estos momentos se puede construir la función de cuasi-verosimilitud (es cuasi-verosimilitud porque los u_t no son normales) y obtener mediante su maximización con algún algoritmo programable la estimación QLM de los parámetros. El filtro también puede aplicarse para la obtención de estimaciones de la volatilidad condicionadas a la información disponible hasta el instante anterior o condicionadas a toda la muestra (estimaciones suavizadas). El problema de aplicar esta técnica es la no normalidad de los errores del modelo en el espacio de los estados, que en modelos asimétricos además están correlados. El filtro de Kalman en ausencia de normalidad proporciona en teoría, estimaciones con error cuadrático medio mínimo entre las estimaciones de linealidad pero no es robusto[24].

El problema de robustez unido a la búsqueda de procedimientos válidos cuando la transformación en el espacio de estados sea no lineal conduce a procedimientos que combinan la simulación de densidades condicionales usando MCMC y estimación Bayesiana o algoritmos EM (estimación-maximización) para la estimación clásica[25]. Debemos destacar como referencias sobre estos métodos de estimación aplicados a modelos SV, los artículos de Shephard(1993) y Kim and Shephard(1994) que implementan un primer algoritmo EM con simulaciones (SIEM), los artículos de Jacquier, Polson and Rossi(1994,95) que realizan una implementación Bayesiana aplicada a modelos SV univariantes y multivariantes respectivamente y las mejoras propuestas de estos métodos aplicando muestreos Gibbs por bloques en los artículos de Shephard(1994b), Shephard and Pitt(1997), Pitt and Shephard(1998) y Kim, Shephard and Chib(1998).

4.4. Otros modelos SV

Extensiones univariantes

Al igual que hacíamos para los modelos ARCH se puede generalizar la estructura de los procesos SV para captar otros hechos estilizados no contemplados por los modelos básicos. Entre las

[24] Ver Peña and Guttman(1989).

[25] Como referencias de estos procedimientos, ver por ejemplo, los trabajos de Carlin, Polson and Stoffer(1992), Carter and Kohn(1994,96) y Frühwirth-Schnatter(1994) y demás referencias en subsección 7.3.

extensiones univariantes estudiadas por la literatura destaquemos el modelo asimétrico ARV(p) de Harvey and Shephard(1993b) definido por:

$$\left.\begin{array}{lcl} \varepsilon_t & = & z_t\sqrt{h_t} \\ \log h_t & = & \gamma_0 + \sum_{i=1}^{p} \gamma_i \log h_{t-i} + \sigma_v v_t \\ \begin{pmatrix} z_t \\ v_t \end{pmatrix} & \sim & \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right) \end{array}\right\}. \tag{4.11}$$

El modelo asimétrico ARV(p) permite modelizar el hecho estilizado del efecto apalancamiento pero tiene la limitación de representar únicamente correlaciones entre rentabilidad y volatilidad de un periodo y no es infrecuente observar correlación entre volatilidad y rentabilidad dos o más periodos hacia atrás. Un modelo alternativo que podría incorporar esos efectos es el propuesto por Peña and Ruiz(1994).

Otras extensiones se obtienen aplicando las ideas habituales en el espacio de estados tomando:

$$\left.\begin{array}{lcl} \varepsilon_t & = & z_t \exp\left(\mathbf{c}_t'\mathbf{h}_t^b/2\right) \\ \mathbf{h}_t^b & = & \mathbf{T}_t\mathbf{h}_{t-1}^b + \boldsymbol{\eta}_t \\ z_t & \sim & \mathrm{IID}(0,1) \\ \boldsymbol{\eta}_t & \sim & \mathrm{IID}(\mathbf{0},\mathbf{H}_t) \end{array}\right\}, \tag{4.12}$$

con $\mathbf{c}_t$ un vector de parámetros, $\mathbf{h}_t^b$ un vector de log-volatilidades ($h_{ti}^b = \log h_{ti}$, $\forall i$) y $\mathbf{T}_t$ una matriz de parámetros. Estos modelos incluyen estructuras ARMA y modelizaciones de efectos estacionales. Un modelo sobre estas ideas es sugerido en Harvey and Shephard(1993c) y Carter and Kohn(1993).

Como alternativa al proceso no estacionario de raíz unitaria, Breidt, Crato and de Lima(1993) y Harvey(1993) proponen para modelizar persistencia un modelo SV fraccionado.

Extensiones multivariantes

La literatura centra su atención en dos grupos de modelos (N = número de series analizadas conjuntamente):

a.-) Modelo ARV(p) multivariante:

$$\left.\begin{array}{lcl} \varepsilon_{ti} & = & z_{ti}\sqrt{h_{ti}},\ i = 1,2,\ldots,N \\ \log \mathbf{h}_t & = & \boldsymbol{\gamma}_0 + \sum_{i=1}^{p} \boldsymbol{\Gamma}_i \log \mathbf{h}_{t-i} + \mathbf{v}_t \\ \mathbf{z}_t & \sim & \mathrm{IID}(\mathbf{0},\boldsymbol{\Sigma}_z) \\ \mathbf{v}_t & \sim & \mathrm{IID}(\mathbf{0},\boldsymbol{\Sigma}_v) \end{array}\right\}, \tag{4.13}$$

donde $\mathbf{z}_t = (z_{t1},\ldots,z_{tN})'$, $\log \mathbf{h}_t = (\log h_{t1},\ldots,\log h_{tN})'$, $\boldsymbol{\gamma}_0$ un N-vector de parámetros, $\boldsymbol{\Gamma}_i$ una matriz $N \times N$ diagonal de parámetros.

b.-) Modelo SV K-factorial:

$$\left.\begin{array}{rcl}\mathbf{y}_t &=& \mathbf{B}\mathbf{f}_t + \varepsilon_t \\ \varepsilon_t &\sim& \text{IID}(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon) \\ \mathbf{f}_t &\sim& \text{proceso SV(p) multivariante}\end{array}\right\}, \tag{4.14}$$

donde $\mathbf{f}_t$ es un K-vector de factores y $\boldsymbol{\Sigma}_\varepsilon$ es la matriz de varianzas-covarianzas de $\boldsymbol{\varepsilon}_t = (\varepsilon_{t1}, \ldots, \varepsilon_{tN})'$.

Harvey, Ruiz and Shephard(1994) estudian el modelo multivariante ARV(1) normal con $\boldsymbol{\Gamma}_1 = \mathbf{I}$ (esto es, el modelo multivariante con raíz unitaria) empleando QLM. Otros trabajos sobre este tipo de modelos son: Mathieu and Schotman(1994) y Jacquier, Polson and Rossi(1995) que emplean un algoritmo EM con simulación basada en MCMC y estimación Bayesiana aplicando MCMC respectivamente.

El modelo SV K-factorial (al que hay que añadirle restricciones para asegurar la identificabilidad de los parámetros) ha sido estudiado en su versión 1-factorial normal con $\boldsymbol{\Sigma}_v$ y $\boldsymbol{\Sigma}_\varepsilon$ diagonales y $p = 1$ por Pitt and Shephard(1995) y Jacquier, Polson and Rossi(1995) empleando métodos de estimación que incluyen simulación por Monte Carlo. Y en una implementación con 3-factores normal más general por Aguilar and West(1998).

Shephard and Pitt(1998) proponen una generalización del modelo K-factorial normal con $\boldsymbol{\Sigma}_v$ diagonal asumiendo que $\boldsymbol{\varepsilon}_t$ sigue asimismo un proceso multivariante ARV(1) normal con matriz de varianzas-covarianzas asociada, $\boldsymbol{\Sigma}_\varepsilon$, asimismo diagonal. La estimación (aplicada al caso $K = 1$) es una estimación Bayesiana basada en simulación por MCMC.

5. Comparación entre los modelos ARCH y SV

Aunque a lo largo de la exposición ya hemos comentado algunas similitudes y diferencias entre las modelizaciones ARCH y SV, es interesante recoger todas ellas en una sección para su reflexión y discursión conjunta. Esta comparación se basa en los artículos de Taylor(1994), Ghysels, Harvey and Renault(1996) y Ruiz(1993,96).

a.-) *En relación a su construcción.* Los modelos ARCH asumen la volatilidad como observable y tienen como única fuente aleatoria para recoger efectos sobre la media y volatilidad a z_t. Los modelos SV parten de la hipótesis conceptualmente más aceptable de volatilidad no observable y separan en dos fuentes aleatorias z_t y v_t los efectos sobre la media y la volatilidad respectivamente.

b.-) *En relación a los hechos estilizados.* Los procesos ARCH y SV normales modelizan adecuadamente hechos estilizados de las series financieras: los rendimientos de los activos son martingalas diferencia, la varianza condicionada cambia en el tiempo y las distribuciones marginales son simétricas y leptocúrticas. Además modelizan los efectos de agrupación de volatilidad y algunas extensiones los efectos de apalancamiento y persistencia de volatilidad.

c.-) *En relación a los modelos de volatilidad continuos aplicados en valoración de activos.* Los modelos SV con retardos y asimétricos son las discretizaciones "exactas" de un amplio grupo de modelos de volatilidad estocástica continuos (Meddahi and Renault(1995)). Por su parte, los modelos ARCH pueden ser vistos como filtros para extraer el proceso de varianza condicional en tiempo continuo a partir de datos financieros discretos (Nelson(1990, 92, 95, 96) y Nelson and Foster(1994, 95)) o como aproximaciones a modelos de valoración continuos Duan(1997).

d.-) *En relación a los parámetros.* La mayor parte de los modelos ARCH necesitan de restricciones iniciales sobre los parámetros para asegurar varianza condicional positiva. Este tipo de restricción no es necesaria en los modelos SV.

e.-) *En relación con los momentos.* Para los modelos GARCH(1,1) y ARV(1) normales tenemos que la distribución marginal de ε_t es simétrica y leptocúrtica. Sin embargo notemos las siguiente diferencias:

- La condición de estacionariedad de los procesos GARCH(1,1) es amplia pero no estricta y en cambio la de los procesos ARV(1) es amplia y estricta a la vez.
- Para asegurar la existencia de momentos de orden superior a 2 en los modelos GARCH(1,1) hay que introducir restricciones adicionales a las de estacionariedad. En cambio, los ARV(1) estacionarios tienen momentos marginales de cualquier orden finitos sin introducir más restricciones.
- El coeficiente de curtosis del modelo GARCH(1,1) depende únicamente de los parámetros α_1 y β_1 y en consecuencia modelizar mayores curtosis supone asumir otras distribuciones condicionales alternativas a la normal. En cambio, en el modelo ARV(1) es función de σ_v^2 (mayor σ_v^2 implica mayor curtosis) con independencia de los valores de los parámetros.
- En relación a las funciones de autocorrelación de ε_t^2 se tiene que para un mismo valor de curtosis: las autocorrelaciones del modelo ARV(1) son menores y tienden a cero más despacio que las correspondientes de un modelo GARCH(1,1). Otra diferencia está en que las autocorrelaciones de ε_t^2 para el modelo GARCH(1,1) son positivas siempre, mientras que en el modelo ARV(1) pueden ser negativas (si $\gamma_1 < 0$)

f.-) *En relación con la propiedad de agregación temporal.* Hay una clase de modelos SV con retardo que son clase cerrada en relación con la propiedad de agregación temporal. Los modelos ARCH, en cambio, no son clase cerrada respecto a la propiedad de agregación temporal.

g.-) *En relación a la estimación.* Los modelos SV son más difíciles de estimar (desde un punto de vista clásico) que los modelos ARCH ya que la función de verosimilitud es de forma implícita en los primeros y explícita en los segundos. Por otra parte, la estimación QLM (basada en el filtro de Kalman) y los modelos de estimación clásicos y Bayesianos que combinan simulación por MCMC de condicionales aplicados en modelos SV permiten obtener estimaciones de la volatilidad dada la información disponible hasta el periodo anterior y también estimaciones suavizadas dada toda la información muestral. Estas estimaciones suavizadas no son posibles (sin realizar aproximaciones) para los modelos ARCH estimados por ML o QML.

h.-) *En relación a evidencias empíricas.* Tenemos evidencias empíricas a favor de los modelos SV frente a modelos ARCH en Hsieh(1991), Danielsson(1994), Kim and Shephard(1994), Jacquier, Polson and Rossi(1994), Ruiz(1996) y Gallant, Hsieh and Tauchen(1997). Y evidencia a favor de los modelos ARCH frente al modelo ARV(1) normal en León and Mora(1996). En cualquier caso, los estudios empíricos no presentan evidencias totalmente decisivas a favor de una de las dos modelizaciones, que por otra parte dan lugar a buenos ajustes.

Como conclusión debemos establecer que los modelos SV aunque han tenido un desarrollo teórico menor y presentan una estimación más compleja constituyen una alternativa válida, interesante y más flexible que la de los modelos ARCH. Debemos indicar también que las otras modelizaciones de series temporales financieras que recogemos en la próxima sección exceptuando los modelos para microestructura de mercados tienen características estructurales más próximas a los modelos SV que a los ARCH.

6. Otras modelizaciones de series temporales financieras

Incluiremos a continuación otros tres grupos de modelos que representan adecuadamente algunas características de series temporales financieras pero que no son modelos típicamente ARCH ni SV.

6.1. Modelos en el espacio de estados

Los modelos SV, como hemos visto en el apartado anterior, presentan características estadísticas interesantes para la modelización de series temporales financieras pero son difíciles de estimar. Shephard(1994a) propone como alternativa al modelo SV, el modelo gaussiano de escala local dado por las siguientes ecuaciones[26]:

$$\left.\begin{array}{rclrcl} \varepsilon_t &=& z_t/\sqrt{h_t^c}, & z_t &\sim& \mathrm{N}(0,1) \\ h_t^c &=& \exp(r_t)h_{t-1}^c\gamma_t, & \gamma_t &\sim& \mathrm{Be}(\delta a_{t-1},(1-\delta)a_{t-1}) \\ h_0^c &\sim& \mathrm{Ga}(a_0,b_0) & & & \end{array}\right\}, \tag{6.1}$$

donde $r_t = \psi(a_{t-1})-\psi(\delta a_{t-1})$, $\psi(.)$ es la función Euler-Psi o digamma, $\delta \in (0,1)$ es un parámetro desconocido y $a_0 > 0$ y $b_0 > 0$ son escalares conocidos.

El modelo gaussiano de escala local supone una modelización en el espacio de los estados con variable de estado la inversa de la volatilidad al cuadrado ($h_t^c = h_t^{-1}$), con la propiedad de filtro

[26]Si definimos $h_t^c = \delta^{-1}h_{t-1}^c\gamma_t$ se tiene que para $\delta < 1$, h_t^c tiende a 0 casi seguramente, la definición alternativa propuesta en (6.1) elimina este problema.

exacto. Efectivamente, siguiendo el artículo de Shephard(1994a) si partimos de $h^c_{t-1}|\mathbf{y}^{(t-1)} \sim \mathrm{Ga}(a_{t-1}, b_{t-1})$ se tiene, aplicando la ecuación sobre la dinámica de h^c_t, que:

$$\begin{aligned} h^c_t|\mathbf{y}^{(t-1)} &\sim \mathrm{Ga}(a_{t|t-1}, b_{t|t-1}), \\ a_{t|t-1} &= \delta a_{t-1}, \\ b_{t|t-1} &= \exp(-r_t) b_{t-1}, \end{aligned} \tag{6.2}$$

e incorporando la nueva observación que:

$$\begin{aligned} h^c_t|\mathbf{y}^{(t)} &\sim \mathrm{Ga}(a_t, b_t), \\ a_t &= a_{t|t-1} + 1/2, \\ b_t &= b_{t|t-1} + \varepsilon_t^2/2. \end{aligned} \tag{6.3}$$

A partir de este filtro (ecuación (6.2) y (6.3)) se obtiene la condicional de $y_t|y^{(t-1)}$:

$$y_t|y^{(t-1)} \sim \mathrm{t}(0, b_{t|t-1} a^{-1}_{t|t-1}, 2a_{t|t-1}). \tag{6.4}$$

A partir de (6.4) y considerada una muestra se construye la función de verosimilitud, y maximizando esta función se obtienen los estimadores máximo verosímiles de los parámetros. Aplicando el filtro también podemos obtener estimadores de la volatilidad condicionados a la información disponible o a toda la muestra (suavizados).

El modelo gaussiano de escala local presenta propiedades estadísticas comunes con el modelo ARV(1) con raíz unitaria y al modelo IGARCH(1,1) y puede generalizarse dentro del espacio de los estados (con lo que permite la introducción de estructura dinámica sobre la media) manteniendo la propiedad de tener filtros exactos. La generalización de estos modelos cuando h^c_t es una matriz de precisiones $\mathbf{H}^c_t$ también es posible manteniendo la propiedad de filtro exacto, consideremos por ejemplo, la siguiente versión simplificada (quitando la estructura dinámica de la media) del modelo de Uhlig(1997):

$$\left.\begin{aligned} \varepsilon_t &= \mathrm{U}((\mathbf{H}^c_t)^{-1})'\mathbf{z}_t, & \mathbf{z}_t &\sim \mathrm{N}(0,1) \\ \mathbf{H}^c_t &= \mathrm{U}(\mathbf{H}^c_{t-1})'\mathbf{\Gamma}_t \mathrm{U}(\mathbf{H}^c_{t-1})/\delta, & \mathbf{\Gamma}_t &\sim \mathrm{Be}_N(a/2, 1/2) \\ \mathbf{H}^c_0 &\sim \mathrm{W}(n_0, \mathbf{S}_0) & & \end{aligned}\right\}, \tag{6.5}$$

donde U(.) es la función que obtiene la matriz triangular superior de la descomposición de Cholesky de una matriz definida positiva, $\mathrm{Be}_N(.)$ es la distribución Beta multivariante generalizada[27] y W(.) representa a la distribución Wishart multivariante, $a > 0$ y $\delta > 0$ son parámetros y $n_0 > 0$ y $\mathbf{S}_0$ son escalar y matriz definida positiva conocidos.

Los modelos (6.1) y (6.5) son la formalización teórica (ver también Uhlig(1992)) de una serie de trabajos basados en el aprendizaje sobre varianza descontada. La idea inicialmente desarrollada por Ameen and Harrison(1985) y Harrison and West(1987) en el caso univariante, y Quintana(1985,87) en el caso matricial, asume en el caso univariante[28], $y_t \sim \mathrm{N}(0, \phi_t^{-1})$ y considera una estructura temporal sobre ϕ_t con el objeto de mantener una estructura conjugada

[27] La definición y propiedades de la distribución Beta multivariante original y generalizada (para unas restricciones más amplias sobre los parámetros) aparecen estudiadas en Muirbead(1982) y Uhlig(1994) respectivamente

[28] La generalización al caso matricial es paralela y será obviada en esta exposición.

y poder obtener filtros exactos., Concretamente, asumen una estructura dinámica tal que si $\phi_{t-1}|\mathbf{y}^{(t-1)} \sim \mathrm{Ga}(n_{t-1}/2, d_{t-1}/2)$ se cumple que $\phi_t|\mathbf{y}^{(t-1)} \sim \mathrm{Ga}(\delta n_{t-1}/2, \delta d_{t-1}/2)$, con $0 < \delta < 1$. La constante δ se conoce como factor de descuento, por representar estocásticamente la evolución de la serie en términos de reducción de la información (o aumento de la dispersión) para ver esta interpretación basta con observar que bajo la estructura anterior: $\mathrm{E}(\phi_{t-1}|\mathbf{y}^{(t-1)}) = \mathrm{E}(\phi_t|\mathbf{y}^{(t-1)})$ y $\mathrm{V}(\phi_t|\mathbf{y}^{(t-1)}) = \delta^{-1}\mathrm{V}(\phi_{t-1}|\mathbf{y}^{(t-1)})$.

Estas técnicas han sido aplicadas en la modelización de series temporales financieras por Putnam and Quintana(1994, 95), Quintana and Putnam(1996) y Quintana, Chopra and Putnam(1995).

6.2. Modelos con cambios de régimen de la volatilidad

Al modelizar la volatilidad de una serie temporal financiera debemos considerar, junto con cambios suaves y progresivos de la misma relacionados con la volatilidad en instantes anteriores posibles cambios en el nivel de la serie[29]. La literatura presenta varias aproximaciones a la modelización de esos cambios, así:

- Engle and Lee(1992) proponen el modelo Component-GARCH(1,1) que supone cambios de nivel de la volatilidad suaves alrededor de un valor de volatilidad fijo de largo plazo (ver (3.16)).

- Otra posibilidad es la de modelizar el número de llegadas de información no habitual, m_t, con una distribución Poisson o Bernoulli y, condicionalmente a m_t asumir que las log-rentabilidades son normales con media $m_t\theta$ y varianza $\sigma_t^2 = \sigma_\epsilon^2 + m_t\sigma_\zeta^2$. Adicionalmente se puede asumir que σ_ϵ^2 sigue un proceso ARCH. Ver, por ejemplo, Jorion(1988) y Vlaar and Palm(1993) y sus referencias.

- Otros autores, por ejemplo Inclán(1993) y sus referencias, consideran que los cambios de volatilidad se producen "por épocas", esto es, durante un intervalo de tiempo (que no sabemos cuando empieza ni termina) la volatilidad tiene un determinado valor constante.

- Otra posibilidad, siguiendo las ideas de Hamilton(1989), es considerar la existencia de una variable aleatoria indicadora del estado no observable que sigue un determinado proceso estocástico, de modo que el valor del estado en t determina el valor de la volatilidad en t.

Nos centraremos a efectos descriptivos, en las dos últimas soluciones y a la modelización de cambios únicamente en la volatilidad asumiendo que no hay estructura dinámica en la media.

El modelo (modificado para no presentar estructura dinámica en media) de Inclán(1993) considera que en el periodo estudiado $[1, f]$, la varianza del proceso cambia N_f veces, el modelo es el siguiente:

[29]En relación a la detección de cambios de nivel de la varianza señalemos los trabajos de Inclán(1993), Inclán and Tiao(1994) y Chen and Gupta(1997).

$$
\left.\begin{array}{rcl}
\varepsilon_t & = & z_t\sqrt{h_t} \\
z_t & \sim & \mathrm{N}(0,1) \\
h_t & = & \left\{\begin{array}{lll} h_1, & \text{si} & 1 \leq t \leq k_1, \\ h_2, & \text{si} & k_1 < t \leq k_2, \\ \vdots & \vdots & \\ h_{N_f}, & \text{si} & k_{N_f} < t \leq f, \end{array}\right.
\end{array}\right\}, \qquad (6.6)
$$

donde los k_i representan los índices de la serie anteriores al cambio, en cada instante del tiempo t hay una probabilidad λ de observar un cambio (esto es, asociado a cada instante de tiempo hay una variable no observable Bernoulli con probabilidad de éxito λ), no son observables k_i ni N_f.

Inclán(1993) emplea este modelo para detectar cambios de varianza sobre la log-rentabilidad de acciones IBM. El modelo (6.6) es conceptualmente sencillo y apropiado para detectar cambios estructurales de volatilidad debidos a ciclos económicos, pero pierde los ajustes suaves de los modelos ARCH y SV y en principio puede dar lugar a "demasiados" cambios de nivel (uno por cada agrupación de volatilidad) que serían difíciles de interpretar en términos económicos y financieros.

Otra alternativa para modelizar cambios de régimen de la volatilidad consiste en considerar modelos con cambios estocásticos en la volatilidad. Consideraríamos, en concreto, modelos tales que:

$$
\left.\begin{array}{rcl}
\varepsilon_t & = & z_t\sqrt{h(S_t)} \\
z_t & \sim & \mathrm{N}(0,1) \\
S_t & \sim & \text{proceso de Markov de orden } d
\end{array}\right\}, \qquad (6.7)
$$

donde S_t es una variable de estado no observable que toma valores en $\{1, 2, \ldots, K\}$ (K-estados) y $h(S_t = i) = h_i$ es la volatilidad asociada al instante t cuando se produce el estado i. Un modelo que se ajusta a estas hipótesis es el modelo de McCulloch and Tsay(1993) para cambios de la varianza para dos estados y $Pr(S_t = 1|S_t = 2) = Pr(S_t = 2|S_t = 1)$. El modelo (6.7) como sucedía con el modelo (6.6) no permite un ajuste parsimonioso de la volatilidad, un modelo alternativo que sí permite este ajuste es el propuesto por Hamilton and Susmel(1994) que consiste en sustituir en (6.4) $z_t \sim \mathrm{N}(0,1)$ por $z_t \sim$ proceso ARCH. Otras modelizaciones recogen cambios en la tendencia, cambios en la varianza en modelos en el espacio de estados, etc., ver por ejemplo Frühwirth-Schnatter(1998) y referencias.

La estimación de los modelos (6.6) y (6.7) y sus posibles generalizaciones para representar cambios suaves unidos a posibles cambios de nivel (por ejemplo, tomar $z_t \sim$ procesos SV, modelizar h_t en (6.6) o $h(S_t)$ en (6.7) tomando un proceso autoregresivo, etc.) puede resultar muy compleja por métodos convencionales clásicos. Como indicación para su implementación mediante técnicas de simulación MCMC señalemos: Shephard(1994), Carter and Kohn(1994, 96) y Frühwirth-Schnatter(1998), este último en una aproximación totalmente Bayesiana.

6.3. Modelización de la microestructura de mercados financieros

En la actualidad los mercados continuos permiten tener observaciones del mercado de alta frecuencia con intervalos de minutos o simplemente por intervalos (obviamente irregulares) entre transacciones. El estudio de estos datos de alta frecuencia debe permitirnos relacionar rendimientos con las características del mercado, por ejemplo comportamiento de los agentes, tipo de agentes, número, volumen de operaciones, etc., entre otros tipos de estudios que también pueden incluir la contrastación de teorías sobre microestructura. Para ver una recopilación sobre las teorías de microestructura el lector interesado puede dirigirse al libro de O'Hara(1995).

En las secciones anteriores hemos comentado que podemos aproximar modelos de valoración continuos mediante el empleo de modelos ARCH y SV. Esta consideración motiva, en principio, la modelización de series temporales financieras de alta frecuencia empleando estos modelos, revisemos por ejemplo el modelo propuesto en Baillie and Bollerslev(1990). Con estos modelos podemos explicar los hechos estilizados habituales de las series temporales financieras pero no podemos representar otras características de interés en microestructura de mercados como son las variaciones de rentabilidad en relación con el tiempo entre transacciones, las variaciones de liquidez intra-día y tiempo entre transacciones, etc. Para recoger estas relaciones se impone establecer un modelo que recoja conjuntamente la variable objeto de observación y el tiempo entre observaciones. Los trabajos de Engle and Russell(1998) y Russell and Engle(1998) presentan algunas propuestas para estos modelos de microestructura.

Como primera fase de la construcción del modelo, Engle(1998) y Engle and Russell(1987) definen el modelo ACD (Autoregressive Conditional Duration) para modelizar el tiempo entre transacciones (duración) en sus versiones exponencial y Weibull. Si denotamos por d_t al tiempo entre las transacciones en los instantes $t-1$ y t, el modelo ACD(m,n) se define por:

$$\left.\begin{array}{rcl} d_t & = & u_t\psi_t \\ u_t & \sim & \mathrm{IID}(1;\phi) \\ \psi_t & = & \omega + \sum_{i=1}^{n}\alpha_i d_{t-i}^2 + \sum_{i=1}^{m}\beta_i\psi_{t-1} \\ & & \omega>0,\ \alpha_i\geq 0,\ \beta_i\geq 0,\ \forall i \end{array}\right\}, \tag{6.8}$$

donde por IID$(1;\phi)$ denotamos una distribución sobre u_t con media 1 y parámetro (o parámetros) ϕ.

Veamos algunas características de este modelo. Notemos que de (6.8), denotando por $\mathbf{d}^{(t-1)} = (d_0,\ldots,d_{t-1})'$, se tiene que:

$$\mathrm{E}(d_t|\mathbf{d}^{(t-1)}) = \psi_t = \omega + \sum_{i=1}^{n}\alpha_i d_{t-i}^2 + \sum_{i=1}^{m}\beta_i\psi_{t-1}, \tag{6.9}$$

por tanto, la media del tiempo entre transacciones es observable en t y es función lineal del tiempo de las m duraciones anteriores (m-memoria condicional) y de sus valores esperados de instantes anteriores (sin característica de memoria limitada). Además, el proceso $\{d_t\}$ se puede expresar como un ARMA con ruido no Gaussiano. Y, si asumimos que $u_t \sim \mathrm{Ex}(1)$, tenemos que la intensidad de ocurrencia de transacciones viene dada por:

$$\lambda(t|N_t, t_1, \ldots, t_{N_t}) = \psi_{N_t-1}^{-1}, \tag{6.10}$$

donde N_t es el número de transacciones hasta t y $t_1, \ldots, t_{N_t}$ los instantes en los que se producen esas transaciones. En consecuencia, para el modelo ACD exponencial la intensidad de ocurrencia para $t > t_{N_t}$ es igual a la inversa del valor esperado de la duración entre las transacciones N_t y $N_t + 1$.

Es de interés, observar que la estructura del modelo (6.8) presenta importantes similitudes con la de un proceso GARCH(m,n). Esta similitud no es sólo formal y permite adaptar para el modelo ACD resultados sobre los modelo GARCH y en particular los relativos a estimación de modelos GARCH no normales. Engle and Russell(1998) realiza un estudio sobre las propiedades de los modelos DLM exponenciales y Weibull y presenta argumentos sobre su estimación empleando algoritmos estándar elaborados para estimar modelos GARCH no normales.

Engle(1998) considera para modelizar conjuntamente la rentabilidad (o log-rentabilidad) de los activos en t, y_t, con la duración en t, d_t. Si descomponemos la distribución conjunta de (y_t, d_t) haciendo:

$$f(y_t, d_t|\mathbf{y}^{(t-1)}, \mathbf{d}^{(t-1)}) = f(y_t|\mathbf{y}^{(t-1)}, \mathbf{d}^{(t)})q(d_t|\mathbf{y}^{(t-1)}, \mathbf{d}^{(t-1)}), \tag{6.11}$$

la idea de Engle(1998) consiste en asumir para $q(d_t|\mathbf{y}^{(t-1)}, \mathbf{d}^{(t-1)})$ un modelo ACD y establecer un modelo para y_t condicionado a las observaciones retardadas y al tiempo entre transacciones presente y retardado. Siguiendo esta idea, Russell and Engle(1998) consideran, para modelizar las observaciones, el modelo ACM (Autoregressive Conditional Multinomial) que asume como característica del activo en estudio, y_t, la diferencia entre los precios en t y $t-1$ y supone K posibles valores (niveles) que modeliza a través del $(K-1)$-vector aleatorio $\mathbf{x}_t$ cuyas componentes x_{ti} son variables aleatorias que toman los valores uno o cero según la observación y_t esté en nivel i o no respectivamente. El modelo ACM(p,q,r) viene dado por:

$$\left.\begin{array}{rcl} \pi_t & = & \mathrm{E}(\mathbf{x}_t|\mathbf{x}^{(t-1)}) \\ p(\pi_t) & = & \boldsymbol{\mu} + \sum_{i=1}^{p} \mathbf{A}_i \mathbf{V}_{t-i}^{-1/2}(\mathbf{x}_{t-i} - \pi_{t-i}) + \sum_{i=1}^{q} \mathbf{B}_i \mathbf{x}_{t-i} + \sum_{i=1}^{r} \mathbf{C}_i p(\boldsymbol{\pi}_{t-i}) + e(\mathbf{d}^t, \mathbf{x}^{(t-1)}) \end{array}\right\}, \tag{6.12}$$

con $p(.)$ una función de $\Re^{K-1}$ a $\Re^{K-1}$ (por ejemplo, la función logit, probit o identidad), $\boldsymbol{\mu}$ es un $(K-1)$-vector de parámetros, $\mathbf{A}$ y $\mathbf{B}$ matrices $(K-1) \times (K-1)$ de parámetros, $\mathbf{V}_t = \mathrm{diag}\{\pi_{it}(1-\pi_{it})\}$ y $e(.)$ una función de $\mathbf{d}^{(t)}$ y $\mathbf{x}^{(t-1)}$ que toma valores en $\Re^{K-1}$ (por ejemplo, $e(\mathbf{d}^{(t)}, \mathbf{x}^{(t-1)}) = \mathbf{g}_1 \log d_t + \mathbf{g}_2 d_t + \mathbf{g}_3(d_t/\psi_t) + \mathbf{g}_4\psi_t$ con $\mathbf{g}_i$ $(K-1)$-vectores de parámetros, $i = 1, \ldots, 4$).

Para su correcta y completa definición, los parámetros y funciones $p(.)$ y $e(.)$ serán tales que $\boldsymbol{\pi}_t$ cumpla las propiedades de una probabilidad.

El artículo de Russell and Engle(1998) realiza una descripción más detallada sobre las propiedades del modelo ACM(p,q,r) "puro" (con $e(.) = 0$), y sobre las restricciones que pueden incorporarse a los parámetros para asegurar que $\boldsymbol{\pi}_t$ sea una probabilidad.

Juntando (6.8) y (6.12) obtenemos el modelo de microestructura ACM(p,q,r)-ACD(m,n) estudiado por Russell and Engle(1998). Notemos para acabar, que el modelo ACD(m,n) puede servir para modelizar conjuntamente (siguiendo (6.11)) otros modelos de microestructura que consideran otras relaciones de interés. Un ejemplo de esta aplicación es el modelo conjunto de liquidez y duración estudiado por Engle and Lange(1997).

7. Estimación Bayesiana de los modelos

La estimación Bayesiana es una alternativa muy interesante frente a la clásica para los modelos que hemos presentado en este trabajo para la modelización de series temporales financieras. Esta afirmación se apoya en los siguiente argumentos:

a.-) La estimación Bayesiana permite introducir de forma adecuada la información inicial del investigador permitiendo además la intervención en cualquier instante de la serie para introducir información adicional no disponible en el instante inicial. En ausencia de información inicial (o cuando no se quiere introducir información), la estimación Bayesiana puede realizarse igualmente usando distribuciones iniciales impropias no informativas o distribuciones propias con mucha dispersión.

b.-) Los procedimientos Bayes están vinculados a la estimación de series temporales financieras a través de la estimación de modelos en el espacio de los estados (el filtro de Kalman puede interpretarse como un proceso inferencial Bayesiano) y de los métodos que incluyen simulación de condicionales por métodos basados en cadenas de Markov de Monte Carlo (MCMC). Y hemos visto que estas dos técnicas son las de aplicación en la estimación de modelos discretos de volatilidad estocástica (SV).

c.-) La estimación Bayesiana obtiene distribuciones posteriores de los parámetros y volatilidad basadas en la información disponible en un instante t o dada toda la muestra (suavizadas) además de las distribuciones predictivas para la característica observada de la serie. Proporciona por tanto, mayor información para la estimación que la estimación clásica.

d.-) La estimación máximo verosímil o cuasi-máximo verosímil de los modelos ARCH no permite (sin hacer uso de filtros aproximativos) la obtención de estimadores de la volatilidad suavizados. La estimación Bayes de estos modelos sí permite la obtención de estos estimadores.

e.-) La estimación máximo verosímil o cuasi-máximo verosímil de los modelos ARCH requiere la maximización de la función de verosimilitud restringida a una serie de condiciones sobre los parámetros para asegurar la positividad de la varianza condicional, estacionariedad y existencia de momentos marginales de orden superior a dos. Estas restricciones complican el problema de estimación considerablemente y sobre todo cuando hay muchos parámetros (por ejemplo, en los modelos multivariantes). En la estimación Bayes estas restricciones se introducen de manera natural al definir las distribuciones iniciales de los parámetros y la existencia de más restricciones o más parámetros no supone en general mayor complicación en el proceso de estimación.

La literatura sobre modelización de series temporales financieras ha captado las ventajas de la estimación Bayesiana y ha producido trabajos de estimación (sobre todo estos últimos 5 años) para procesos ARCH y SV. Las próximas dos subsecciones analizan las contribuciones más interesantes clasificadas en tres grupos: aportaciones en modelos ARCH, en modelos SV y otras aportaciones de interés.

7.1. Aportaciones Bayes a la estimación de modelos ARCH

Los trabajos pioneros en la estimación Bayesiana de los modelos ARCH se deben a Geweke(1988a,b, 89a,b) y se refieren a la estimación de procesos ARCH(q). Estos trabajos están motivados por las razones (d) y (e) que comentábamos al principio de la sección, consideran distribuciones iniciales impropias para los parámetros, y la estimación Bayes hace uso de la integración por Monte Carlo. Otros dos trabajos de interés, que hacen uso de la integración por Monte Carlo para modelos GARCH(p,q) son Kleibergen and Van Dijk(1993) y Koop(1994). Dos trabajos recientes de Müller and Pole(1997) y Bauwens and Lubrano(1998) proponen estimación Bayesiana de modelos GARCH(p,q) haciendo uso de simulación MCMC, en particular Gibbs sampler. Revisamos a continuación estas aportaciones con mayor detalle.

Geweke(1988a) propone la estimación Bayes del modelo de regresión ARCH con proceso asociado ARCH(q) normal, empleando tres distribuciones iniciales impropias dadas por:

$$\begin{aligned} \pi_a(\boldsymbol{\theta}) &\propto \left[(1-\textstyle\sum_{i=1}^q \alpha_i)/\omega\right]^{1/2} J(\boldsymbol{\theta})K(\boldsymbol{\theta}), \\ \pi_b(\boldsymbol{\theta}) &\propto J(\boldsymbol{\theta})K(\boldsymbol{\theta}), \\ \pi_c(\boldsymbol{\theta}) &\propto J(\boldsymbol{\theta}), \end{aligned} \tag{7.1}$$

donde $\boldsymbol{\theta} = (\mathrm{b}', \boldsymbol{\phi}')'$, b es el vector con los parámetros de la regresión, $\boldsymbol{\phi}$ un vector que recoge los parámetros del proceso ARCH(q), $J(\boldsymbol{\theta})$ es una función indicatriz que toma el valor 1 si $\omega > 0$ y $\alpha_i \geq 0$, $i = 1, \ldots, q$ (condiciones para que la varianza sea positiva) y $K(\boldsymbol{\theta})$ una función indicatriz que toma el valor 1 si $\sum_{i=1}^q \alpha_i < 1$ (condición de estacionariedad). El procedimiento de trabajo propuesto consiste en la simulación de $\boldsymbol{\theta}$ aplicando muestreo importante y la estimación por integración por Monte Carlo de la media (cuando existe) de una función de los parámetros $l(\boldsymbol{\theta})$ de interés en el problema (la convergencia asintótica del estimador a la normal y del estimador varianza a la varianza asintótica aparece estudiada en Geweke(1989b)). Como detalles particulares de la implementación señalemos la elección de la función importante. Respecto a este punto, Geweke(1988a) indica que el empleo de las distribuciones asintóticas para **b** y $\boldsymbol{\phi}$ como función importante produce resultados pobres y sugiere emplear para $\mathbf{b}|\boldsymbol{\phi}$ la distribución asintótica normal y para $\boldsymbol{\phi}$ una modificación de la distribución asintótica $\mathrm{t}(\widehat{\boldsymbol{\phi}}, \mathbf{V}, \nu)$ en la que $\mathbf{V}$ se ajusta a lo largo de cada eje y en cada dirección para reducir el cociente $\frac{L(\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{I(\boldsymbol{\theta})}$ con $I(\boldsymbol{\theta})$ la función importante. Más detalles en Geweke(1989b).

Esta implementación se aplicó sobre un proceso simulado y sobre datos reales sobre la inflacción U.S. en el periodo 1951-79 (el ejemplo real también aparece referenciado y discutido en Geweke(1988b)) obteniendo gráficos de las distribuciones posteriores de los parámetros y la varianza no condicionada.

Geweke(1989a) considera nuevamente el modelo anterior con distribución inicial sobre los parámetros impropia $\pi_a(\boldsymbol{\theta})$ (definida en (7.1)) y estudia la obtención de las predictivas, de la media de una función de los valores futuros y los factores Bayes para producir tests entre varios modelos ARCH con distinto valor q. Como detalles de la implementación señalemos que se sugiere el empleo de aceleración antitética en la integración de Monte Carlo y que la simulación de $f(y_{T+1},\dots,y_{T+s}|y_1,\dots,y_T,\boldsymbol{\theta})$ (con T el tamaño de la muestra) se realiza aplicando recursivamente la ecuación $y_t = \mathbf{x}_t'\mathbf{b} + \varepsilon_t$ para $t = T + j$ con $j = 1,\dots,s$ (s el número de pasos hacia el futuro en la predicción). En esta ocasión se implementa este modelo para los precios diarios de acciones IBM, empleándose los cálculos de los factores Bayes para decidir el valor de q del modelo ARCH que mejor ajusta a los datos.

La metodología de trabajo basada en muestreo importante ha sido aplicada dentro del análisis de modelos GARCH por otros autores para distribuciones condicionales de los errores no normales. Así, el trabajo de Kleibergen and Van Dijk(1993) al realizar un análisis estadístico Bayes sobre estacionariedad de los modelos GARCH, considera errores condicionales t y obtiene simulaciones de las distribuciones posteriores de los parámetros por muestreo importante y los primeros momentos de algunas funciones de los parámetros de interés en el análisis usando integración por Monte Carlo. Y el trabajo de Koop(1994) que considera una distribución sobre los errores condicionados no paramétrica asumiendo que $z_t = \varepsilon_t/\sqrt{h_t} \sim f(z_t) = (\sum_{i=1}^{m} q_j z_t^j)\Phi(z_t)$, con $\Phi(.)$ la función de densidad de una normal estándar. En los dos trabajos junto a la estructura ARCH del modelo se considera una estructura dinámica sobre la media de las observaciones (ver (2.1)).

Un enfoque de trabajo alternativo y de mayor utilidad cuando el número de parámetros es grande para realizar un análisis Bayes en los modelos ARCH se basa en la implementación de cadenas de Markov Monte Carlo, en particular Gibbs sampler, para simular las distribuciones posteriores y predictivas. Al implementar esta técnica el problema se presenta en la construcción de las distribuciones condicionadas que tienen que ser habitualmente aproximadas. Dentro de esta línea metodológica podemos situar los trabajos de Müller and Pole(1997) y Bauwens and Lubrano(1998).

El artículo de Müller and Pole(1997) tiene como objetivo principal la estimación de un modelo GARCH multivariante dinámico aplicando cadenas de Markov pero presenta como resultado previo el estudio del modelo de regresión con errores modelizados según un proceso GARCH(1,1) normal que exponemos a continuación. Si denotamos por: $\boldsymbol{\theta} = (\mathbf{b}', \omega, \alpha_1, \beta_1)' = (\mathbf{b}', \boldsymbol{\phi}')'$, por $\mathbf{b}$ el vector de parámetros de la regresión, por $g_\phi(\boldsymbol{\phi}|\mathbf{b})$ y $g_b(\mathbf{b}|\boldsymbol{\phi})$ dos distribuciones de prueba para la simulación de $\boldsymbol{\phi}$ y $\mathbf{b}$ respectivamente, y por $f(\boldsymbol{\theta}|y_1,\dots,y_T)$ a la distribución final de $\boldsymbol{\theta}$, el algoritmo de Metropolis propuesto por Müller and Pole(1997), para obtener las distribuciones finales de $\boldsymbol{\phi}$ y $\mathbf{b}$, realiza los siguientes pasos para pasar de $\boldsymbol{\theta}^t$ a $\boldsymbol{\theta}^{t+1}$ y $\boldsymbol{\theta}^{t+2}$:

Algoritmo 7.1:

a.-) Simula $\tilde{\boldsymbol{\phi}} \sim g_\phi(\boldsymbol{\phi}|\mathbf{b})$ y considera el candidato $\tilde{\boldsymbol{\theta}} = (\mathbf{b}^{t\prime}, \tilde{\boldsymbol{\phi}}')'$.

b.-) El candidato $\tilde{\boldsymbol{\theta}}$ se acepta como nuevo estado con probabilidad:

$$a(\boldsymbol{\theta}^t, \tilde{\boldsymbol{\theta}}) = \min\left(1, \frac{f(\tilde{\boldsymbol{\theta}}|y_1, \ldots, y_T)}{g_\phi(\tilde{\boldsymbol{\phi}}|\mathbf{b}^t)} \frac{g_\phi(\boldsymbol{\phi}^t|\mathbf{b}^t)}{f(\boldsymbol{\theta}^t|y_1, \ldots, y_T)}\right).$$

Si el candidato es aceptado $\boldsymbol{\theta}^{t+1} = \tilde{\boldsymbol{\theta}}$ y si no lo es $\boldsymbol{\theta}^{t+1} = \boldsymbol{\theta}^t$.

c.-) Simula $\tilde{\mathbf{b}} \sim g_b(\mathbf{b}|\boldsymbol{\phi}^{t+1})$ y considera el candidato $\tilde{\boldsymbol{\theta}} = (\tilde{\mathbf{b}}', \boldsymbol{\phi}^{t+1\prime})'$.

d.-) El candidato $\tilde{\boldsymbol{\theta}}$ se acepta con probabilidad $a(\boldsymbol{\theta}^t, \tilde{\boldsymbol{\theta}})$. Si hay aceptación $\boldsymbol{\theta}^{t+2} = \tilde{\boldsymbol{\theta}}$ y si no la hay $\boldsymbol{\theta}^{t+2} = \boldsymbol{\theta}^{t+1}$.

Para evitar el problema de rechazar demasiado, se introducen algunos pasos de Metropolis cada 10 iteraciones de la cadena anterior (ver artículo para detalles sobre esta subcadena). Las distribuciones de prueba se obtienen de la siguiente manera:

- Obtención de $g_b(\mathbf{b}|\boldsymbol{\phi})$. Sea $\widehat{\mathbf{b}}$ el valor actual de $\mathbf{b}$, computamos a partir de la ecuación de regresión el valor de ε_{t-1} y a partir de éste y h_{t-1} se obtiene h_t. La distribución posterior de $\mathbf{b}$ en este modelo auxiliar es la distribución de prueba.

- Obtención de $g_\phi(\boldsymbol{\phi}|\mathbf{b})$. Se deriva de manera similar a partir del modelo: $\varepsilon_t^2 = \omega + \alpha_1\varepsilon_{t-1}^2 + \beta_1 h_{t-1} + u_t$, con $u_t \sim N(0, \sigma^2)$ y $p(\sigma) \propto 1/\sigma$.

Este procedimiento de trabajo es aplicado para estudiar los logaritmos de las diferencias de tipo de cambio diario dolar/marco alemán. Las variables explicativas en el modelo de regresión son: un término constante, los logaritmos de las diferencias dolar/marco del periodo anterior y los logaritmos de las diferencias de cambio dolar/franco suizo y dolar/yen también del periodo anterior.

7.2. Aportaciones Bayes a la estimación de modelos SV

La estimación Bayes del modelo SV está muy motivada por las dificultades que tiene la estimación de estos modelos y por los avances en estimación Bayes de modelos en el espacio de los estados (ya comentábamos que las técnicas de estimación QLM de estos modelos se basan en la escritura de los mismos en el espacio de estados). La primera contribución en estimación Bayes que hace uso de de simulación por cadenas de Monte Carlo es el artículo de Jacquier, Polson and Rossi(1994). Estos autores sugieren un algoritmo Gibbs sampler de movimiento sencillo (single-move algorithm) en el que las volatilidades h_t son simuladas una a una condicionadas a $\mathbf{h}_{\backslash t}$ (denotamos así a un vector que contiene todos los h_s, $s = 1, \ldots, T$ con $s \neq t$ y T el tamaño de la muestra), los parámetros y las observaciones. El procedimiento de estimación de Jacquier, Polson and Rossi(1994) es empíricamente más eficiente que las técnicas GMM y QML clásicas pero la convergencia cuando los parámetros γ son próximos a uno es todavía lenta. Una razón para esta lentitud es la correlación entre las volatilidades condicionadas a la muestra. El trabajo de Shephard and Pitt(1997), recogiendo las soluciones propuestas por la literatura de estimación de modelos en el espacio de los estados, propone un algoritmo Gibbs sampler de movimiento conjunto o por bloques (multimove algorithm). Los artículos de Jacquier, Polson and Rossi(1994) y Shephard and Pitt(1997) se refieren a la estimación del modelo ARV(1) normal (ver (4.1)). Pero son extensibles a otros modelos SV, en particular, Jacquier, Polson and Rossi(1995) estiman un SV multivariante y Shephard and Pitt(1998) aplica el algoritmo de Shephard and Pitt(1997) al análisis de un modelo SV factorial multivariante.

Para una revisión de estos trabajos, consideremos el modelo ARV(1) (ver (4.1)) con distribución condicional normal, denotemos por $\boldsymbol{\theta} = (\gamma_0, \gamma_1, \sigma_v)'$ y asumamos siguiendo a Jacquier, Polson and Rossi(1994) la siguiente distribución inicial sobre $\boldsymbol{\theta}$:

$$\left.\begin{array}{lcl} \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix} & \sim & \mathrm{N}\left(\begin{pmatrix} \overline{\gamma_0} \\ \overline{\gamma_1} \end{pmatrix}, \sigma_v^2 \mathbf{A}^{-1}\right) \\ \sigma_v & \sim & \mathrm{Ga}^{-1/2}(\nu_0, s_0^2) \end{array}\right\}, \tag{7.2}$$

donde $\mathbf{A}$ es una matriz definida positiva conocida, $\overline{\gamma_0}$ y $\overline{\gamma_1}$ son escalares conocidos y ν_0 y s_0^2 son escalares positivos conocidos.

Dada una muestra de T observaciones, $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_T)'$, Jacquier, Polson and Rossi(1994) proponen el siguiente algoritmo Gibbs[30] de movimiento sencillo para simular las distribuciones posteriores $\mathbf{h}|\boldsymbol{\varepsilon}$ y $\boldsymbol{\theta}|\boldsymbol{\varepsilon}$ con $\mathbf{h} = (h_1, \dots, h_T)'$. El algoritmo pasa de $(\mathbf{h}^{t\prime}, \boldsymbol{\theta}^{t\prime})$ a $(\mathbf{h}^{t+1\prime}, \boldsymbol{\theta}^{t+1\prime})$ de la siguiente manera:

Algoritmo 7.2:

a.-) Se simulan recursivamente $h_i^{t+1} \sim f(h_i|\mathbf{h}_{\backslash i}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon})$ para $i = 1, \dots, T$.

b.-) Se obtienen simulaciones de $(\gamma_0^{t+1}, \gamma_1^{t+1})' \sim f((\gamma_0, \gamma_1)'|\mathbf{h}^{t+1}, \sigma_v^t, \boldsymbol{\varepsilon})$ y de $\sigma_v^{t+1} \sim f(\sigma_v|\mathbf{h}^{t+1}, \gamma_0^{t+1}, \gamma_1^{t+1}, \boldsymbol{\varepsilon})$

donde $\mathbf{h}_{\backslash i}^t$ es el vector $\mathbf{h}^t$ eliminando la componente h_i.

De la forma de las distribuciones iniciales se obtiene que las distribuciones en (b) del algoritmo 7.2 son conocidas normal y gamma inversa ($\mathrm{Ga}^{-1/2}(.)$) y por tanto sencillas de simular. En cambio las distribuciones en (a) son conocidas salvo constante de integración y al no tener una forma distribucional conocida son más difíciles de simular. En concreto, se tiene que para $i = 1, \dots, T$:

$$\begin{aligned} f(h_i|\mathbf{h}_{\backslash i}, \boldsymbol{\theta}, \boldsymbol{\varepsilon}) &= f(h_i|h_{i-1}, h_{i+1}, \boldsymbol{\theta}, \varepsilon_i) \\ &\propto h_i^{-0.5} \exp(-0.5\varepsilon_i^2/h_i) 1/h_i \exp(-(\log h_i - \mu_i)^2/(2\sigma^2)), \end{aligned} \tag{7.3}$$

donde $\mu_i = (\gamma_0(1-\gamma_1) + \gamma_1(\log h_{i+1} + \log h_{i-1}))/(1-\gamma_1^2)$ y $\sigma^2 = \sigma_v^2/(1+\gamma_1^2)$.

Para simular (7.3) podemos aplicar varios procedimientos. Así, Geweke(1994) en su comentario al artículo de Jacquier, Polson and Rossi(1994) señala que la distribución (7.3) es log-concava

[30] Geweke(1996) hace una revisión sobre los métodos de integración por Monte Carlo y simulación por MCMC, presentando un ejemplo con el modelo ARV(1).

y sugiere simularla empleando el algoritmo de Wild and Gilks(1993). Otra posibilidad es la de aproximar (7.3) a través de una distribución conocida (y por tanto fácil de simular) $g(h_i|\mathbf{h}_{\backslash i}, \boldsymbol{\theta}, \boldsymbol{\varepsilon})$ de modo que:

$$f(h_i|\mathbf{h}_{\backslash i}, \boldsymbol{\theta}, \boldsymbol{\varepsilon}) \leq cg(h_i|\mathbf{h}_{\backslash i}, \boldsymbol{\theta}, \boldsymbol{\varepsilon}), \tag{7.4}$$

e introducir dentro del algoritmo Gibbs un algoritmo de Metropolis-Hasting para su simulación. La elección de la función $g(h_i|\mathbf{h}_{\backslash i}, \boldsymbol{\theta}, \boldsymbol{\varepsilon})$ y en qué medida es buena aproximación de $f(h_i|\mathbf{h}_{\backslash i}, \boldsymbol{\theta}, \boldsymbol{\varepsilon})$ determina una menor o mayor número de rechazos y por lo tanto un menor o mayor número de iteraciones del algoritmo Gibbs 7.2 combinado con el algoritmo Metropolis-Hastings.

En relación a la construcción de $g(h_i|\mathbf{h}_{\backslash i}, \boldsymbol{\theta}, \boldsymbol{\varepsilon})$, Jacquier, Polson and Rossi(1994) observan que el primer término de (7.3) es una $\mathrm{Ga}^{-1}(0.5, 0.5\varepsilon_i^2)$ y aproximan el término log-normal mediante una $\mathrm{Ga}^{-1}(.)$ con media y varianza las del término log-normal. Proponen, por tanto, la distribución $g(h_i|\mathbf{h}_{\backslash i}, \boldsymbol{\theta}, \boldsymbol{\varepsilon}) = \mathrm{Ga}^{-1}(a_i, b_i)$ con $a_i = 0.5 + (1 - 2\exp\sigma^2)/(1 - \exp\sigma^2)$ y $b_i = 0.5\varepsilon_t^2 + (a_t - 1)\exp(\mu_t + 0.5\sigma^2)$. Y eligen c computando la mediana de los radios $f(h_i|\mathbf{h}_{\backslash i}, \boldsymbol{\theta}, \boldsymbol{\varepsilon})/g(h_i|\mathbf{h}_{\backslash i}, \boldsymbol{\theta}, \boldsymbol{\varepsilon})$ en tres puntos centrados en la moda de la distribución $g(h_i|\mathbf{h}_{\backslash i}, \boldsymbol{\theta}, \boldsymbol{\varepsilon})$. Elegida la función aproximativa y c señalan que la relación de dominancia (7.4) no se cumple para todo punto por lo que proponen el siguiente algoritmo Metropolis-Hastings a incluir en el algoritmo 7.2, para pasar de h_i^t a h_i^{t+1}, $i = 1, \ldots, T$:

Algoritmo 7.3:

a.-) Simulo $\tilde{h}_i \sim g(h_i|\mathbf{h}_{\backslash i}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon})$.

b.-) El candidato se acepta con probabilidad:

$$a(h_i^t, \tilde{h}_i) = \min\left(1, \frac{f(\tilde{h}_i|\mathbf{h}_{\backslash i}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon})\ \min\{f(h_i^t|\mathbf{h}_{\backslash i}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon}), cg(h_i^t|\mathbf{h}_{\backslash t}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon})\}}{f(h_i^t|\mathbf{h}_{\backslash i}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon})\ \min\{f(\tilde{h}_i|\mathbf{h}_{\backslash i}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon}), cg(\tilde{h}_t|\mathbf{h}_{\backslash t}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon})\}}\right).$$

Si el candidato es aceptado $h_i^{t+1} = \tilde{h}_i$, en otro caso $h_i^{t+1} = h_i^t$.

Pitt and Shephard(1995) proponen otra aproximación, en este caso normal, para la simulación de $h_i^b = \log h_i$, $i = 1, \ldots, T$. De (7.3) tenemos que:

$$\begin{aligned} f(h_i^b|\mathbf{h}_{\backslash i}^b, \boldsymbol{\theta}, \boldsymbol{\varepsilon}) &= f(h_i^b|h_{i-1}^b, h_{i+1}^b, \boldsymbol{\theta}, \varepsilon_i) \\ &\propto\ -0.5h_i^b \exp(-0.5\varepsilon_i^2 \exp(-h_i^b)) \exp(-(h_i^b - \mu_i)^2/(2\sigma^2)), \end{aligned} \tag{7.5}$$

y tomando logaritmos y usando el desarrollo de Taylor alrededor de μ_i de primer orden que:

$$\begin{aligned} &\log f(h_i^b|\mathbf{h}_{\backslash i}^b, \boldsymbol{\theta}, \boldsymbol{\varepsilon}) \\ \leq&\ \text{const.} - 0.5h_i^b - (h_i^b - \mu_i)^2/(2\sigma^2) - 0.5\varepsilon_i^2\{(1 + \mu_i)\exp(-\mu_i) - h_i^b \exp(-\mu_i)\} \\ =&\ \text{const.'} + \log g(h_i^b|\mathbf{h}_{\backslash i}^b, \boldsymbol{\theta}, \boldsymbol{\varepsilon}). \end{aligned} \tag{7.6}$$

Pitt and Shephard(1995) a partir de (7.6) proponen emplear como función aproximada de $f(h_i^b|\mathbf{h}_{\setminus i}^b, \boldsymbol{\theta}, \boldsymbol{\varepsilon})$ la distribución $g(h_i^b|\mathbf{h}_{\setminus i}^b, \boldsymbol{\theta}, \boldsymbol{\varepsilon}) = N(h_i^b|\mu_i^*, \sigma^2)$ con $\mu_i^* = \mu_i + 0.5\sigma^2[\varepsilon_i^2 \exp(-\mu_i) - 1]$. De (7.6) se deduce que se cumple la propiedad (7.4) para todo punto y en consecuencia el algoritmo 7.2 Gibbs resultante de sustituir h_i por h_i^b (*algoritmo 7.2bis*) para $t = 1, \ldots, T$ se completa con el siguiente algoritmo Metropolis-Hastings más sencillo, para pasar de h_i^{bt} a h_i^{bt+1}:

Algoritmo 7.4:

a.-) Simulo $\tilde{h}_t^b \sim g(h_i^b|\mathbf{h}_{\setminus i}^{b\ t}, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon})$.

b.-) El candidato se acepta con probabilidad:

$$a(h_i^{bt}, \tilde{h}_i^b) = \frac{f(\tilde{h}_i^b|\mathbf{h}_{\setminus i}^{b\ t}, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon})}{g(\tilde{h}_i^b|\mathbf{h}_{\setminus i}^{b\ t}, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon})}.$$

Si el candidato es aceptado $h_i^{bt+1} = \tilde{h}_i{}^b$, en otro caso $h_i^{bt+1} = h_i^{bt}$.

Los resultados empíricos de Pitt and Shephard(1995) señalan que su propuesta (algoritmo 7.2bis combinado con el algoritmo 7.4) es más eficiente que la propuesta de Jacquier, Polson and Rossi(1994) (algoritmo 7.2 combinado con el algoritmo 7.3).

La convergencia de la cadena es todavía mejorable si se considera un algoritmo Gibbs sampler de movimiento conjunto o por bloques. Esta posibilidad es recogida en el artículo de Shephard and Pitt(1997) que considera para el modelo ARV(1) normal el siguiente algoritmo Gibbs por bloques, para pasar de $(\mathbf{h}^{t\prime}, \boldsymbol{\theta}^{t\prime})$ a $(\mathbf{h}^{t+1\prime}, \boldsymbol{\theta}^{t+1\prime})$

Algoritmo 7.5:

a.-) Se simulan conjuntamente $\mathbf{h}_{t_{k_i},k_i}^{t+1} \sim f(\mathbf{h}_{t_{k_i},k_i}|h_{t_{k_i}-1}^t, h_{t_{k_i}+k_i+1}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon}_{t_{k_i},k_i})$ para $i = 1, \ldots, K$, donde $\mathbf{h}_{t,k} = (h_i, \ldots, h_{t+k})$, $\boldsymbol{\varepsilon}_{t,k} = (\varepsilon_t, \ldots, \varepsilon_{t+k})$, los k_i son longitudes aleatorias de los K bloques en los que se dividen los tiempos $\{1, \ldots, T\}$ y t_{k_i} representa el instante de tiempo asociado a la primera variable del bloque i-ésimo. ($k_i = \text{int}[T((i + U_i)/(K+2))]$, $U_i \sim \text{Un}(0,1)$, $i = 1, \ldots, K$.)

b.-) Se obtienen simulaciones de $(\gamma_0^{t+1}, \gamma_1^{t+1})' \sim f((\gamma_0, \gamma_1)'|\mathbf{h}^{t+1}, \sigma_v^t, \boldsymbol{\varepsilon})$ y de $\sigma_v^{t+1} \sim f(\sigma_v|\mathbf{h}^{t+1}, \gamma_0^{t+1}, \gamma_1^{t+1}, \boldsymbol{\varepsilon})$

Analogamente al algoritmo 7.2, la parte de implementación más compleja corresponde a la simulación de los bloques $\mathbf{h}_{t_{k_i},k_i}^{t+1} \sim f(\mathbf{h}_{t_{k_i},k_i}|h_{t_{k_i-1}}^t, h_{t_{k_i+k_i+1}}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon}_{t_{k_i},k_i})$ para $i = 1, \ldots, K$. Shephard and Pitt(1997), consideran la simulación por bloques de los $h_i^b = \log h_i$, $i = 1, \ldots, T$ (en el correspondiente *algoritmo 7.5bis*) y proponen la "extensión" de los argumentos planteados en Pitt and Shephard(1995) al caso multivariante. La idea básica consiste en aproximar $\log f(\mathbf{h}_{t_k,k}^b|h_{t_k-1}^b{}^t, h_{t_k+k+1}^b{}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon}_{t_k,k})$ usando el desarrollo por Taylor hasta segundas derivadas a una expresión const + $\log g(\mathbf{h}_{t_k,k}^b|h_{t_k-1}^b{}^t, h_{t_k+k+1}^b{}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon}_{t_k,k})$ con $g(\mathbf{h}_{t_k,k}^b|h_{t_k-1}^b{}^t, h_{t_k+k+1}^b{}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon}_{t_k,k})$ una distribución normal multivariante. Y en la implementación de un algoritmo de Metropolis-Hastings para la simulación de $f(\mathbf{h}_{t_k,k}^b|h_{t_k-1}^b{}^t, h_{t_k+k+1}^b{}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon}_{t_k,k})$ a partir de su aproximación

$g(\mathbf{h}^b_{t_k,k}|h^b_{t_k-1}{}^t, h^b_{t_k+k+1}{}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon}_{t_k,k})$ (que en este caso no necesariamente domina). Para la simulación de la distribución $g(\mathbf{h}^b_{t_k,k}|h^b_{t_k-1}{}^t, h^b_{t_k+k+1}{}^t, \boldsymbol{\theta}^t, \boldsymbol{\varepsilon}_{t_k,k})$, Shephard and Pitt(1997) sugiere aprovechar la forma de la aproximación que puede interpretarse dentro de los modelos en el espacio de estados y aplicar el algoritmo de de Jong and Shephard(1995).

7.3. Otras aportaciones de interés

Ya hemos visto que un grupo interesante de modelos para series temporales financieras se basa en los modelos en el espacio de estados o tiene su interpretación transformada en el mismo. Por esta razón es interesante tener algunas referencias sobre la literatura Bayes sobre la estimación de estos modelos, citemos, el libro de West and Harrison(1997) y dentro de la estimación Bayes que hace uso de algoritmos MCMC, los artículos de Carlin, Polson and Stoffer(1992), Frühwirth-Schnatter(1994), Shephard(1994b), Carter and Kohn(1994, 96), y los algoritmos de suavizado de de Jong and Shephard(1995) y Knorr-Held(1998).

Sobre la estimación Bayes de modelos con cambios de régimen señalemos Frühwirth-Schnatter (1998) y sus referencias.

8. Conclusiones

Para concluir destaquemos que a pesar del gran trabajo realizado en la modelización de series temporales financieras que se desprende de la recopilación presentada, aún queda mucho trabajo por hacer en esta área tanto en la modelización como en la estimación y en la diagnosis del modelo. A modo de ejemplo, algunas líneas de investigación pueden ser:

- La formulación y estimación de modelos más generales que presenten estructura dinámica sobre la media y volatilidad.
- La estimación de modelos SV más generales.
- El desarrollo de tests para contrastar entre modelos cuando éstos requieren para su estimación de simulaciones por MCMC.
- La generalización de modelos de cambio de régimen de volatilidad que contemplen cambios progresivos de la misma y/o estructura dinámica para la media.
- El análisis de los modelos de microestructura presentados y formulación de nuevos modelos.

Por último, referirnos a los modelos no lineales: bilineales, autoregresivos exponenciales, etc. (ver por ejemplo, Tong(1990)). Algunos modelos no lineales (p.e. los bilineales) cumplen la propiedad de correlación marginal nula de los residuos y presencia de correlación marginal en los cuadrados que presentan las series financieras por lo que en principio pueden ser candidatos válidos para su modelización. La no inclusión en este trabajo se debe a que estos modelos no tienen una clara interpretación financiera. Así, si analizamos, por ejemplo, los primeros momentos condicionales de los modelos bilineales detectamos estructura dinámica para la media y una varianza condicional constante, lo que contrasta con la idea desarrollada por los modelos de esta recopilación (en concordancia con las teorías financieras de valoración) de modelización dinámica de la volatilidad condicional. Las diferencias entre los modelos bilineales y ARCH, los test entre los dos tipos de modelizaciones y los test para detectar la presencia conjunta de una estructura para los residuos bilineal y una modelización de las volatilidades ARCH han sido estudiadas por la literatura ver, por ejemplo, Hsieh(1989) y Higgins and Bera(1991).

Referencias

Aguilar, O. and West, M. (1998). Bayesian dynamic factor models and variance matrix discounting for portfolio allocation. Discussion paper. Duke University.

Ameen, J.R. and Harrison, P.J. (1995). Normal discount Bayesian models. *Bayesian Statistics 2.* (Bernardo, DeGroot, Lindley and Smith, eds.) North Holland. Amsterdam and Valencia University Press.

Andersen, T.G. (1994). Stochastic autoregressive volatility: a framework for volatility modeling. *Math. Finance.* **4**, 75-102.

Andersen, T.G. and Sorensen, B.E. (1996). GMM estimation of a stochastic volatility model: a Monte Carlo study. *J. Business Econom. Statist.* **14**(3), 328-352.

Baba, Y., Engle, R.F., Kraft, D.F. and Kroner, K.F. (1990). Multivariate simultaneous generalized ARCH. *Mimeo.* Department of Economics, University of California at San Diego.

Baillie, R.T. and Bollerslev, T. (1989). The message in daily exchange rates: a conditional variance tale. *J. of Business and Economic Statistics.* **7**, 297-305.

Baillie, R.T. and Bollerslev, T. (1990). Intra-day and inter-market volatility in foreing exchange rates. *Review of Economic Studies.* **58**, 565-585.

Ballie, R.T. and Bollerslev, T. (1992). Prediction in dynamic models with time-dependent conditional variances. *J. Econometrics.* **52**, 91-113.

Baillie, R.T., Bollerslev, T. and Mikkelsen, O. (1993). Fractionally integrated generalized autoregressive conditional heteroskedasticity. Working paper. Michigan State University.

Bates, D.S. (1995). Testing option pricing models. *Handbook of Statistics* (Maddala and Rao, eds). **14**, 567-611.

Bauwens, L. and Lubrano, M. (1998). Bayesian inference on GARCH models using the Gibbs sampler. *The Econometrics Journal.* Forthcoming.

Bera, A.K. and Higgins, M.L. (1993). ARCH models: properties, estimation and testing. *J. of Economic Surveys.* **7**(4), 305-366.

Berndt, E.K., Hall, B.H., Hall, R.E. and Hausman, J.A. (1974). Estimation and inference in nonlinear structural models. *Annals of Economic and Social Measurement.* **3**, 653-665.

Black, F. (1976). Studies in stock price volatility changes. *Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section, Amer. Statist. Assoc.* 177-181.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *J. of Political Economy.* **81**, 637-659.

Bollerslev, T. (1986). A generalized autoregressive conditional heteroskedasticity. *J. of Econometrics.* **31**, 307-327.

Bollerslev, T. (1987). A conditional heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics.* **69**, 542-547.

Bollerslev, T. (1988). On the correlation structure of the generalized autoregressive conditional heteroskedastic process. *J. of Time Series Analysis.* **9**, 121-131.

Bollerslev, T. (1990). Modeling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH approach. *Review of Economics and Statistics.* **72**, 498-505.

Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992). ARCH modeling in finance. A review of the theory and empirical evidence. *J. of Econometrics.* **52**, 5-59.

Bollerslev, T., Engle, R.F. and Wooldridge, J.M. (1988). A capital asset pricing model with time-varying covariances. *J. of Political Economy.* **96**, 116-131.

Bollerslev, T., Engle, R.F. and Nelson, D.B. (1995). ARCH models. *The Handbook of Econometrics.* (R.F. Engle and D. McFadden eds). **4**. Forthcoming.

Bollerslev, T. and Wooldridge, J.M. (1992). Quasi maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Rev.* **11**, 143-172.

Bougerol, P. and Picard, N. (1992). Stationarity of GARCH process and some nonnegative time series. *J. of Econometrics.* **52**, 115-128.

Box, G.E.P. and Tiao, G.C. (1973). *Bayesian inference in statistical analysis.* John Willey. New York.

Breidt. F.J. and Carriquiri, A.L. (1996). Improved quasi-maximum likelihood estimation for stochastic volatility models. *Modeling and prediction. Honoring Seymour Geisser.* (Lee, Johnson and Zellner, eds). Springer-Verlag. New York.

Breidt, F.J., Crato, N. and de Lima, P. (1993). Modeling long-memory stochastic volatility. Discussion paper. Iowa State University.

Carlin, B.P., Polson, N.G. and Stoffer, D.S. (1992). A Monte Carlo approach to nonnormal and nonlinear state-space modeling. *J. of the American Statistics Association.* **87**, 493-500.

Carter, C.K. and Kohn, R. (1993). On the applications of Markov chain Monte Carlo methods to linear state space models. *Proceedings of the ASA, Business and Economic Statistics Section*, 131-136.

Carter, C.K. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika.* **81**(3), 541-553.

Carter, C.K. and Kohn, R. (1996). Markov chain Monte Carlo in conditionally Gaussian state space models. *Biometrika.* **83**(3), 589-601.

Christie, A.A. (1982). The stochastic behavior of common stock variances: value, leverage, and interest rate effects. *J. of Financial Economics.* **10**, 407-432.

Clarke, P.K. (1973). A subordinated stochastic process model with fixed variance for speculative prices. *Econometrica.* **41**, 135-156.

Crowder, M.J. (1976). Maximum likelihood estimation with dependent observations. *J. Roy. Statist. Soc. Ser. B.* **38**, 45-53.

Danielsson, J. (1994). Stochastic volatility in asset prices: estimation with simulated maximum likelihood. *J. Econometrics.* **61**, 375-400.

De Jong, P. and Shephard, N. (1995). The simulation smoother for time series models. *Biometrika.* **82**(2), 339-350.

Diebold, R.X. and Nerlove, M. (1989). The dynamics of exchange rate volatility: a multivariate latent factor ARCH model. *J. of Applied Econometrics.* **4**, 1-21.

Duan, J.C. (1995). The GARCH option pricing model. *Math. Finance.* **5**, 13-32.

Duan, J.C. (1997). Augmented GARCH (p,q) process and its diffusion limit. *J. of Econometrics.* **79**, 97-127.

Drost, F.C. and Nijman, T.E. (1992). Temporal aggregation of GARCH processes. *Econometrica.* **61**, 909-927.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflaction. *Econometrica.* **50**, 987-1008.

Engle, R.F. (1987). Multivariate ARCH with factor structures – cointegration in variance. Discussion paper 87-27, University of California at San Diego.

Engle, R.F. (1998). The econometrics of ultra-higt frequency data. Unpublished paper, Department of Economics, University of California at San Diego.

Engle, R.F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews.* **5**, 1-87.

Engle, R.F. and González-Rivera, G. (1991). Semiparametric ARCH models. *J. of Business and Economic Statistics.* **9**, 345-359.

Engle, R.F., Granger, C.W.J. and Kraft, D. (1986). Combining competitive forecast of inflaction using a bivariate ARCH model. *J. of Economic Dynamics and Control.* **8**, 151-165.

Engle, R.F. and Kraft, D.F. (1983). Multiperiod forecast error variances of inflaction estimated from ARCH models. *Applied Time Series Analysis of Economic Data* (Zellner ed.) Bureau of the Census, Washington D.C., 293-302.

Engle, R.F. and Lange, J. (1997). Measuring, forecasting and explaining time varying liquidity in the stock market. Discussion paper 97-12, Department of Economics, University of California at San Diego.

Engle, R.F. and Lee, G.G.J. (1992). A permanent and transitory component model of stock return volatility. Unpublished paper: Department of Economics, University of California at San Diego.

Engle, R.F., Lilien, D.M. and Robins, R.P. (1987). Estimating time varying risk premia in the term structure: the ARCH-M model. *Econometrica.* **55**, 391-407.

Engle, R.F. and Ng, V. (1993). Measuring and testing the impact of news on volatility. *J. of Finance.* **48**, 1749-1778.

Engle, R.F., Ng, V. and Rothschild, M. (1990). Asset pricing with a Factor-ARCH covariance structure: empirical estimates for treasury bills. *J. of Econometrics.* **45**, 213-237.

Engle, R.F. and Russell, J.R. (1998). Autoregressive Conditional Duration: a new model for irregularly spaced transaction data. *Econometrica.* Forthcoming.

Fama, E.F. (1963). Mandelbrot and the stable Paretian distribution. *J. Business.* **36**, 420-429.

Fama, E.F. (1965). The behavior of stock market prices. *J. Business.* **38**, 34-105.

Fornari, F. and Mele, A. (1997). Sign- and volatility-switching ARCH models: theory and applications to international stock markets. *J. Applied Econometrics.* **12**, 49-65.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *J. of Time Series Analysis.* **15**(2), 183-202.

Frühwirth-Schnatter, S. (1998). MCMC estimation of classical and dynamic switching and mixture models. Unpublished paper, Department of Statistics, Vienna University of Economics and Business Administration.

Gallant, A.R., Hsieh, D. and Tauchen, G. (1997). Estimation of stochastic volatility models with diagnostics. *J. of Econometrics.* **81**, 159-192.

Geweke, J. (1986). Modelling the persistence of conditional variances: comment. *Econometric Reviews.* **5**, 57-61.

Geweke, J. (1988a). Exact inference in models with autoregressive conditional heteroskedasticity. *Dynamic econometric modeling* (Berndt, White and Barnett, eds). Cambidge University Press. 73-104.

Geweke, J. (1988b). Comments on Poirier: operational Bayesian methods in econometrics. *J. of Economic Perspectives.* **2**, 159-166.

Geweke, J. (1989a). Exact predictive densities for linear models with ARCH disturbances. *J. Econometrics.* **40**, 63-86.

Geweke, J. (1989b). Bayesian inference in econometric models using Monte Carlo integration. *Econometica.* **57**, 1317-1339.

Geweke, J. (1994). Bayesian comparison of econometric models. Federal Reserve Bank of Minneapolis. Working Paper.

Geweke, J. (1996). Monte Carlo simulation and numerical integration. *Handbook of Computational Economics, Volume I* (Amman, Kendrick and Rust, eds.) Elsevier. Amsterdam.

Glosten, L., Jagannathan, R. and Runkle, D. (1993). Relationship between the expected value and the volatility of the nominal excess return on stocks. *J. of Finance.* **48**, 1779-1801.

Gouriéroux, C. (1997). *ARCH models and financial applications.* Springer. New York.

Gysels, E., Harvey, A.C. and Renault, E. (1996). Stochastic volatility. *Handbook of Statistics* (Maddala and Rao, eds). **14**, 119-191.

Hamilton, J.D. (1989). Analysis of time series subject to changes in regime. *J. Econometrics.* **64**, 307-333.

Hamilton, J.D. and Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *J. of Econometrics.* **64**, 307-333.

Hansen, L.P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica.* **50**, 1029-1054.

Harrison, P.J. and West, M. (1986). Bayesian forecasting in practice. *Bayesian Statistics study year.* Report 13. University of Warwick.

Harrison, P.J. and West, M. (1987). Practical Bayesian forecasting. *The Statistician.* **36**, 115-125.

Harvey, A.C. (1981). *The econometric analysis of time series.* Oxford.

Harvey, A.C. (1989). *Forecasting structural time series models and the Kalman filter.* Cambridge.

Harvey, A.C. (1993). Long memory in stochastic volatility. Discussion paper. London School of Economics.

Harvey, A.C., Ruiz, E. and Sentana, E. (1992). Unobserved component time series models with ARCH disturbances. *J. of Econometrics.* **52**, 129-157.

Harvey, A.C., Ruiz, E. and Shephard, N. (1994). Multivariate stochastic variance models. *Rev. Econom. Stud.* **61**, 247-264.

Harvey, A.C. and Shephard, N. (1993a). Estimation and testing of stochastic variance models. *STICERD Econometrics.* Discussion paper EM93/268. London School of Economics.

Harvey, A.C. and Shephard, N. (1993b). Estimation of an asymmetric stochastic volatility model for asset returns. Discussion paper. Statistics Department. London School of Economics.

Harvey, A.C. and Shephard, N. (1993c). The econometrics of stochastic volatility LSE Financial Markets Group. Discussion paper 166. London School of Economics.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *J. of Finance.* **46**(1), 179-207.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies.* **6**, 327-343.

Higgins, M.L. and Bera, A.K. (1991). ARCH and bilinearity as competing models for nonlinear dependence. *Mimeo.* Department of Economics . University of Wisconsin-Milwaukee.

Higgins, M.L. and Bera, A.K. (1992). A class of nonlinear ARCH models. *International Economic Review.* **33**, 137-158.

Hobson, D.G. (1997). Stochastic volatility. *Statistics in finance* (Hand and Jacka. eds). 283-306.

Hsieh, D. (1989). Testing for nonlinear dependence in daily foreign exchange rate changes. *J. of Business.* **62**, 339-368.

Hsieh, D. (1991). Chaos and nonlinear dynamics: applications to financial markets. *J. of Finance.* **46**, 1836-1877.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *J. of Finance.* **42**(2), 281-300.

Inclán, C. (1993). Detection of multiple changes of variance using posterior odds. *J. Business and Economic Statistics.* **11**(3), 289-300.

Jacquier, E., Polson, N.G. and Rossi, P.E. (1994). Bayesian analysis of stochastic volatility models (with discussion). *J. of Business and Economic Statistics.* **12**, 371-417.

Jacquier, E., Polson, N.G. and Rossi, P.E. (1995). Models and prior distributions for multivariate stochastic volatility. Discussion paper. Graduate School of Business. University of Chicago.

Jorion, P. (1988). On jump processes in the foreign exchange and stock markets. *Review of Financial Studies.* **1**, 427-445.

Kim, S. and Shephard, N. (1994). Stochastic volatility: optimal likelihood inference and comparison with ARCH model. Discussion paper. Nuffield College. Oxford.

Kim, S., Shephard, N. and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review Economic Studies.* **65**. Forthcoming.

King, M., Sentana, E. and Wadhwani, S. (1994). Volatility links between national stock markets. *Econometrica.* **62**, 901-933.

Kleibergen, K. and Van Dijk, H.K. (1993). Non-stacionarity in GARCH models: a Bayesian analysis. *J. Applied Econometrics.* **8**, S41-S61.

Knorr-Held, L. (1998). Conditional prior proposals in dynamic models. *Scandinavian Journal of Statistics.* Forthcoming.

Koop, G. (1994). Bayesian semi-nonparametric ARCH models. *Review of Economics and Statistics.* **76**, 176-181.

Kraft, D.F. and Engle, R.E. (1982). Autoregressive conditional heteroscedasticity in multiple time series models. Unpublished paper, Department of Economics, University of California at San Diego.

Lee, S. and Hansen, B. (1994). Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory.* **10**, 29-52.

León, A. and Mora, J. (1996). Modelling conditional heteroskedasticity: application to stock return index IBEX-35. Working paper. WP-AD 96-11. Instituto Valenciano de Investigaciones Económicas (IVIE).

Lin, W.L. (1992). Alternative estimators for factor GARCH models – A Monte Carlo comparison. *J. Applied Econometrics.* **7**, 259-279.

Lumsdaine, R. (1992). Asymtotic properties of the quasi-maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models: A Monte Carlo investigation. *J. Business Econom. Statist.* **13**, 1-10.

Lumsdaine, R. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationarity GARCH(1,1) models. *Econometrica.* **64**, 575-596.

Mahieu, R. and Schotman, P. (1994). Stochastic volatility and the distribution of exchange rate news. Unpublished paper, Department of Economics, Princeton University.

Mandelbrot, B.B. (1963). The variation of certain speculative prices. *J. Business.* **36**, 394-416.

Mandelbrot, B.B. (1967). The variation of some other speculative prices. *J. Business.* **40**, 363-413.

McCulloch, R.E. and Tsay, R.S. (1993). Bayesian inference and prediction for mean and variance shifts in Autoregressive time series. *J. of the American Statistic Association.* **88**, 968-978.

Meddahi, N. and Renault, E. (1995). Aggregations and marginalisations of GARCH and stochastic volatility models. Discussion paper. GREMAQ.

Melino, A. and Turnbull, S.M. (1990). Pricing foreign currency options with stochastic volatility. *J. of Econometrics.* **45**, 239-265.

Müller, P. and Pole, A. (1997). Monte Carlo posterior integration in GARCH models. Discussion paper. Duke University.

Muirhead, R.J. (1982). *Aspects of multivariate statistical theory.* John Wiley. New York.

Nelson, D.B. (1988). *Time series behavior of stock market volatility and returns.* Ph. D. dissertation. MIT.

Nelson, D.B. (1990a). ARCH models as diffusion approximations. *J. Econometrics.* **45**, 7-39.

Nelson, D.B. (1990b). Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory.* **6**, 318-334.

Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica.* **59**, 347-370.

Nelson, D.B. (1992). Filtering and forecasting with misspecified ARCH models I: getting the right variance with the wrong model. *J. Econometrics.* **25**, 61-90.

Nelson, D.B. (1995). Asymptotic smoothing theory for ARCH models. *J. Econometrics* to appear.

Nelson, D.B. (1996). Asymptotic filtering theory for multivariate ARCH models. *J. Econometrics.* **71**, 1-47.

Nelson, D.B. and Cao, C.Q. (1992). Inequality constraints in the univariate GARCH model. *J. of Business and Economic Statistics.* **10**, 229-235.

Nelson, D.B. and Foster, D.P. (1994). Asymptotic filtering theory for univariate ARCH models. *Econometrica.* **62**, 1-41.

Nelson, D.B. and Foster, D.P. (1995). Filtering and forecasting with misspecified ARCH models II: making the right forecast with the wrong model. *J. Econometrics.* **67**, 303-335.

Ng, V.K., Engle, R.F. and Rothschild, M. (1992). A multi-dynamic-factor model for stock returns. *J. of Econometrics.* **52**, 245-266.

Nijman, T.E. and Sentana, E. (1996). Marginalization and contemporaneous aggregation in multivariate GARCH processes. *J. Econometrics.* **71**, 71-81.

O'Hara, M. (1995). *Market microstructure theory.* Blackwell Publishers. Cambridge.

Pagan, A. and Schwert, G.W. (1990). Alternative models for conditional stock volatility. *J. of Econometrics.* **45**, 267-290.

Palm, F.C. (1996). GARCH models of volatility. *Handbook of Statistics* (Maddala and Rao, eds). **14**, 209-240.

Pantula, S.G. (1985). Estimation of autoregressive models with ARCH errors. Unpublished manuscript. Department of Statistics. North Carolina State University. Raleigh, NC.

Peña, D. and Guttman, I. (1989). Optimal collapsing of mixture distributions in robust recursive estimation. *Communications in Statistics.* **18**(3), 817-833.

Peña, J.I. and Ruiz, E. (1994). Stock market regulations and international financial integration: the case of Spain. Manuscrito. Dpto. Economia de la Empresa. Universidad Carlos III de Madrid.

Pitt, M.K. and Shephard, N. (1995). Parameter-driven exponential family models. Unpublished paper, Nuffield College, Oxford.

Pitt, M.K. and Shephard, N. (1998). Analytic convergence rates and parameterisation issues for the Gibbs sampler applied to state space models. *J. Time Series Analysis.* **19**. Forthcoming.

Putnam, B.H. and Quintana, J.M. (1994). New Bayesian statistical approaches to estimating and evaluating models of exchange rates determination. *Proceedings of the ASA, Section on Bayesian Statistical Science, 1994 Joint Statistical Meetings.* American Statistical Association.

Putnam, B.H. and Quintana, J.M. (1995). Shrink wrap for currencies. *Asset allocation: applying quantitative discipline to asset allocation.* (B.H. Putnam, ed). Global Investor, Euromoney Publications. London. 139-146.

Quintana, J.M. (1985). A dynamic linear matrix-variate regression model. ReseARCH report 83, Department of Statistics, University of Warwick.

Quintana, J.M. (1987). *Multivariate Bayesian forecasting models.* Unpublished Ph.D. thesis, University of Warwick.

Quintana, J.M., Chopra, V.K. and Putnam, B.H. (1995). Global asset allocation: stretching returns by shrinking forecast. *Proceedings of the ASA Section on Bayesian Statistical Science, 1995 Joint Statistical Meetings.* American Statistical Association.

Quintana, J.M. and Putnam, B.H. (1996). Debating currency markets efficiency using multiple-factor models. *Proceedings of the ASA, Section on Bayesian Statistical Science, 1996 Joint Statistical Meetings.* American Statistical Association.

Ruiz, E. (1993). Stochastic volatility versus autoregressive conditional heteroscedasticity. Working paper 93-44. Statistics and Econometrics Series 26. Universidad Carlos III. Madrid.

Ruiz, E. (1994a). Modelos para series temporales heterocedásticas. *Cuadernos Económicos de I.C.E.*. **56**, 73-108.

Ruiz, E. (1994b). Quasi-maximum likelihood estimation of stochastic volatility models. *J. of Econometrics.* **63**, 289-306.

Ruiz, E. (1996). Modeling volatility: comparison between ARCH and stochastic volatility models. Unpublished paper: Departamento de Estadística y Econometría, Universidad Carlos III, Madrid, Spain.

Russell, J.R. and Engle, R.F. (1998). Econometric analysis of discrete-valued irregularly-spaced financial transactions data using a new Autoregressive Conditional Multinomial Model. Unpublished paper, Graduate School of Business, University of Chicago.

Scott, L. (1987). Option pricing when the variance changes randomly: theory, estimation and an application. *J. of Financial and Quantitative Analysis.* **22**, 419-438.

Shephard, N. (1993). Fitting nonlinear time-series models with applications to stochastic variance models. *J. of Applied Econometrics.* **8**, S135-S152.

Shephard, N. (1994a). Local scale models. State space alternative to integrated GARCH processes. *J. of Econometrics.* **60**, 181-202.

Shephard, N. (1994b). Partial non-Gaussian state space. *Biometrika.* **81**(1), 115-131.

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. *Time series models. In econometrics finance and other fields* (Cox, Hinkley and Barndorff-Nielsen, eds). 1-67.

Shephard, N. and Pitt, M.K. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika.* **84**(3), 653-667.

Shephard, N. and Pitt, M.K. (1998). Analysis of time varying covariances: a factor stochastic volatility approach. *Bayesian Statistics.* **6**. (J.M. Bernardo, J.O. Berger, A.P. David and A.F.M. Smith, eds.). Oxford Universtity Press. Forthcoming.

Stein, E.M. and Stein, J. (1991). Stock price distributions with stochastic volatility: an analytic approach. *Review of Financial Studies.* **4**, 727-752.

Taylor, S.J. (1986). *Modeling financial time series.* John Wiley. Chichester.

Taylor, S.J. (1994). Modeling stochastic volatility: a review and comparative study. *Math. Finance.* **4**(2), 183-204.

Tong, H. (1990). *Non-linear time series. A dynamical system approach.* Clarenton Press. Oxford.

Uhlig, H. (1991). BVARTEC- Bayesian vector auto regressions with time varying error-covariances. Discussion paper. Princeton University.

Uhlig, H. (1994). On singular Wishart and singular multivariate Beta distributions. *The Annals of Statistics.* **22**, 395-405.

Uhlig, H. (1997). Bayesian vector autoregressions with stochastic volatility. *Econometrica.* **65**(1), 59-73.

Vlaar, P.J.G. and Palm, F.C. (1993). The message in weekly exchange rates in the European Monetary System: mean reversion, conditional heteroscedasticity and jumps. *J. of Business and Economic Statistics.* **11**(3), 351-360.

Wei, S.X. (1998). A Censored-GARCH model of asset returns with price limits. CORE discussion paper 9815. Université Catolique de Louvain.

Weiss, A.A. (1984). ARMA models with ARCH errors. *J. of Time Series Analysis.* **5**(2) , 129-143.

Weiss, A.A. (1986). Asymptotic theory for ARCH models: estimation and testing. *Econometric Theory.* **2**, 107-131.

West, M. and Harrison, P.J. (1997). *Bayesian forecasting and dynamic models.* 2^{nd} edition. Springer. New York.

Wild, P and Gilks, W.R. (1993). Adaptive rejection sampling from log-concave density functions. *J. of the Royal Statistical Society.* Serie C. **42**, 701-708.

Zakoian, J.M. (1994). Threshold heteroskedastic models. *J. of Economic Dynamics and Control.* **18**, 931-955.